TECHNISCHE UNIVERSITÄT DRESDEN

CEPIE Working Paper No. 05/18

Center of Public and International Economics

STOCHASTIC FRONTIER ANALYSIS IN HIGHER EDUCATION: A SYSTEMATIC REVIEW

December 2018

Sabine Gralka

Editors: Faculty of Business and Economics, Technische Universität Dresden.

This paper is published on the Open Access Repository Qucosa.
The complete Working Paper Series can be found at the CEPIE Homepage | EconStor | RePEc

ISSN 2510-1196

Stochastic Frontier Analysis in Higher Education: A Systematic Review*

Sabine Gralka†
TU Dresden

December 2018

Abstract

This paper provides a systematic review of the literature that employs stochastic frontier analysis to measure the efficiency of higher education institutions. The overview opens with a look at the general development of the literature, before emphasis is laid on the methodical aspects. Focus is thereby placed on the necessary underlying assumptions and the employed specifications, discussing their advantages and drawbacks. Afterwards, the factors that were specified in the literature, including the input and output variables, as well as the determinants of efficiency, are discussed in detail. Based on the insights of the literature review, the paper highlights some of the existing deficiencies and ways forward. To our knowledge, the present study provides the first systematic review on the usage of the stochastic frontier analysis to measure efficiency in the higher education sector.

JEL classification: C14, D61, I22, I23, H52

Keywords: Review, Efficiency, Higher Education, Stochastic Frontier Analysis

---

*Acknowledgments: I am grateful for the valuable comments from Alexander Kemnitz, Patrick Zwerschke, Jasmina Berbegal Mirabent and Dolors Gil-Doménech. Moreover, I thank Charlotte Junge and Madeleine Brühl for their valuable research assistance. This work was partly financed through the project line "Quantitative Wissenschaftsforschung", which is funded by the German Federal Ministry of Education and Research.

†Corresponding author: Technische Universität Dresden, Faculty of Business and Economics, 01062 Dresden, Germany. Email: sabine.gralka@tu-dresden.de

1. Introduction

The topic of the efficiency of universities is increasingly being discussed in the public debate. By now, policy makers frequently rely on efficiency comparisons to monitor and evaluate higher education [HE] institutions, often to subsequently decide on the allocation of funding. Responding to this demand, the literature concerning efficiency evaluations has evolved exponentially during the last decades. However, as varying approaches to measure efficiency have been explored, slowly but surely, it has become difficult to see the wood for the trees. By taking a step back and summarizing where we stand and what we know, this study aims to provide the reader with a structured overview and highlight methodical and practical steps forward. To this end, the present paper reviews empirical studies using stochastic frontier analysis [SFA] to evaluate the HE sector at the institutional level. Starting with the first study published in 1998, we examine all explorations that are published in peer-reviewed journals during the past three decades until and inclusive of 2017. The 63 studies reviewed include 208 different estimations, which vary in the employed underlying sample, the estimation specifications and the factors considered. Based on the systematic review we will answer the following questions: (1) In general, how has the literature on SFA in the HE sector developed over the last three decades? (2) Which SFA specifications were employed to measure the efficiency of academic institutions and what are their advantages? (3) What are the factors, including the inputs and outputs as well as the determinants of efficiency that are used in the frontier literature? In doing so, our research seeks to complement existing reviews on efficiency, such as the survey by De Witte and López-Torres (2017) on the education sector and by Rhaiem (2017) on the HE sector. The deliberate focus of the present study on SFA applications only, allows us to compare the studies in more depth than before, focusing in particular on the substantial methodical development. Aiming to provide a compact overview, we refrain from explaining the reasoning of each author behind each choice. Instead, we emphasize the general advantages of each specification or factor. To our knowledge, the present study provides the first systematic review on the usage of SFA to measure efficiency in the HE sector. By showing what has been done, where to look and what is still to do, the study provides researchers with a profound guide for future work.

We show that there seems to be a core model for the evaluation of efficiency. Most authors use panel data to assess the public universities of their respective countries. The researchers thereby choose to estimate time-variant efficiency with an emphasis on the determinants of efficiency or they aim to take the heterogeneity of institutions into account. While the input of institutions is most commonly represented by the expenditures of universities, teaching is represented by students or graduates and research by the amount of research grants. Nevertheless, there seems to be a lack of consensus in particular regarding the employed determinants of efficiency and the presentation of the results. Rewarding future extensions of the efficiency analysis of HE institution will include cross-country evaluations, the distinction between short- and long-term efficiency, the inclusion of third-mission activities and the evaluation of the dynamics of efficiency.

In the next section, a general introduction is given, including a definition of the term efficiency, a look at its role in the evaluation of HE institutions and a glance at the estimation method. In the following section, the sample selection and coding procedure are discussed, and then the results of the review are given in section four. The review thereby opens with a look at the descriptive statistics, answering question (1). Emphasis is then laid on the methodical approaches that have been used, answering question (2). Afterwards, the main input and output variables specified in the literature are discussed, followed by the portrayal of the determinants of efficiency, answering question (3). Complementing the review, section five discusses some shortcomings of the existing literature and the present paper. Finally, section six presents our conclusions and offers some suggestions for future research.

2. Framework

Within the economic literature, efficiency commonly refers to the evaluation of the used input relative to the obtained output of an institution. An organization is classified as efficient if it produces the largest possible output from a given set of inputs (output-oriented measure) or if it uses the least possible input for a given output (input-oriented measure). To evaluate what the largest possible output (or lowest input) is, a benchmark is needed. One, therefore, assesses a group of firms and classifies the most productive units of the sample as efficient. Hence, the resulting individual efficiency values are always relative measures. Two standard approaches to estimate efficiency exist, namely, Data Envelopment Analysis [DEA], which is a non-parametric method, and the SFA.[1] The latter method, which originated from the study of Aigner, Lovell, and Schmidt (1977), assumes that an underlying mathematical function (commonly a production or cost function) represents the benchmark, the so-called frontier. Institutions that are close to the frontier possess the "best" input to output ratio and are therefore called efficient. Depending on the proximity of the unit to the frontier, each organization obtains a value between zero and one, with a higher value indicating a higher efficiency. The SFA as a parametric method is based on an estimation, which allows the consideration of random errors (noise). The resulting regression table permits an additional check of the quality of the estimation and allows to compare different specifications. However, the necessity of an underlying function for the estimation also restricts

[1] Since the DEA is not the focus of the present study, it is not discussed further. For a comprehensive overview, see Coelli et al. (2005).

the quantity of considered factors. It is possible to either look at one output to multiple inputs or vice versa.

Being one of the most prominent methods to estimate efficiency, the SFA is applied in a variety of research areas, with agriculture and banking being the most prominent (Fried, Lovell and Schmidt, 2008). From those fields, the method also found its way into the evaluation of the HE sector and, given the growing number of publications, surveys on the topic emerged. To our knowledge, Worthington (2001) was the first to review papers that measure the efficiency of educational institutions, which were up to this point mainly DEA studies. Subsequently, with a chosen focus on the DEA method only, Johnes (2006) as well as Emrouznejad, Parker and Tavares (2008) describe recent advancements that have been used for measuring efficiency in education, emphasizing their advantages and drawbacks. A related, but more general discussion on how operational research has been applied to education can be found in Johnes (2015). The two most recent surveys on efficiency evaluations are provided by De Witte and López-Torres (2017), looking at the whole education sector, and Rhaiem (2017), focusing on the HE sector only. Both studies include DEA as well as SFA evaluations in their reviews but, given the higher percentage of studies using the DEA method (77% and 83%), they focus their discussion on it. We, therefore, argue that the present study fills a void in the literature, by focusing on studies using SFA to estimate efficiency in education. This allows the reviewed studies to be compared in more depth than before, looking especially at the substantial methodical development. Moreover, given their different methodical restrictions, the compositions of the chosen input and output factors differ greatly between the two methods. The combined evaluation of DEA and SFA studies can therefore even be viewed as critical to some extent.

3. Scope of Review and Dataset

Our data compilation followed established methods for a systematic review and was conducted in four steps, following Booth, Sutton and Papaioannou (2016). Based on the research question, we established the criteria of inclusion and exclusion (step 1), identified relevant studies (step 2), coded (step 3) and then evaluated them (step 4).

3.1 Criteria of inclusion and exclusion

Five criteria of inclusion were used to select potential studies. To be included in our systematic review, a study had to deal with efficiency in the HE sector, use an SFA for the measurement, be an empirical study (theoretical and conceptual studies were not included) and, due to pragmatic reasons, be written in English and published before or in the year 2017.[2] In addition, we had to impose three criteria for exclusion after assembling the data. First, we have chosen to consider only peer-reviewed journal articles. Consequently, other publication forms, such as books, conference proceedings, newspaper articles, unpublished works, etc., were not considered. While this enables us to ensure the quality of the reviewed studies (to some extent), we realize that it could lead to a slightly biased sample as well as an oversight of recent developments, which are not yet published in journals. We, therefore, included footnotes at the appropriate sections which refer to excluded, but relevant, studies and current developments. The second restriction concerns the aim of the evaluated studies. While SFA is primarily used to measure efficiency, an adjacent literature stream employs the method to look at economies of scale and scope of HE institutions. Since those studies are not necessarily aiming for an accurate measurement of the efficiency term, which is the main methodical development of the SFA in recent years (discussed in section 4.2), we only include studies which explicitly measure efficiency. Studies that measure solely economies of scale and scope are excluded. Finally, we limited our sample to studies examining the efficiency of institutions (universities, faculties, departments, etc.), to obtain a homogenous sample and to be able to compare the utilized factors. Contrary to other surveys, we are not excluding studies based on the ranking of the journals or on the considered country.

3.2 Identification of relevant studies

Our search strategy involved four separate search activities, namely, (1) the search for the most relevant keywords ("efficiency", "frontier", "higher education", "SFA", etc.), (2) back-referencing (reviewing the references of the included studies), (3) citation tracking (reviewing the references in which the included study has been cited) as well as (4) screening and hand-searching in selected electronic journals. The studies were then evaluated according to the above listed criteria in a double-blind review of studies based on careful reading of the titles, keywords and abstracts.

3.3 Coding of selected studies

Following the selection, the sample was coded following a precise coding procedure. For this purpose, a codebook has been developed that specifies the characteristics of the study itself, the employed data, the factors used, the statistical techniques and the study results. Coding decisions were recorded for each study in separate files. Finally, all final coding was double-checked. Since some studies include more than one estimation, if, for example, efficiency is assessed using varying specifications within a given article, we distinguish between studies and estimations in the following evaluation. While the examination of studies provides a general overview, the

[2] Two studies were accepted and made available online by the respective journals in 2017 but were assigned an issue only in 2018. Therefore, the studies by Barra, Lagravinese and Zotti (2018) and Gralka (2018) are evaluated as if they were published 2017 but are referred to as published in 2018.

consideration of estimations allows a more detailed analysis. Nevertheless, the review of the latter can lead to a biased impression when the distribution of estimations over studies is highly skewed.

3.4 Evaluation

Based on the above noted criteria for inclusion, the identification of studies resulted in a total of 129 publications using SFA for the evaluation of HE institutions. Of those studies, 5 are parts of books and 28 are other publication forms, primarily working papers and dissertations. From the remaining 96 journal publications, 52 studies evaluate efficiency, 26 evaluate economies of scale and scope and 18 examine both, resulting in a sample of 70 studies. A further six studies had to be excluded due to the level of observation, and one study was excluded due to missing information.[3] The working sample therefore comprises 63 studies with 208 estimations. Accordingly, a study includes three estimations on average. The outliers worth mentioning are the analysis by Zhang, Bao and Sun (2016), who conduct 11, and Laureti, Secondi and Biggeri (2014), who compare 16 estimations. An overview of the considered studies can be found in Table A1. Each section concludes with a short summary.

4. Review

The following survey starts with a look at the general characteristics of the study (4.1), answering the first research question. Afterwards, the underlying methodical assumptions and the applied SFA specifications are discussed in detail (4.2 and 4.3), responding to the second research question. The third research question is then answered by examining the input and output factors used (4.4) as well as the determinants of efficiency (4.5). We deliberately refrain from reporting efficiency levels from the reviewed studies and avoid statements such as “results in high [or low] efficiency values”. A comparison of efficiency values among studies is not meaningful, due to the relative nature of the values. We recommend the books by Kumbhakar and Lovell (2003) and Coelli et al. (2005) for further reading on the methodical details.

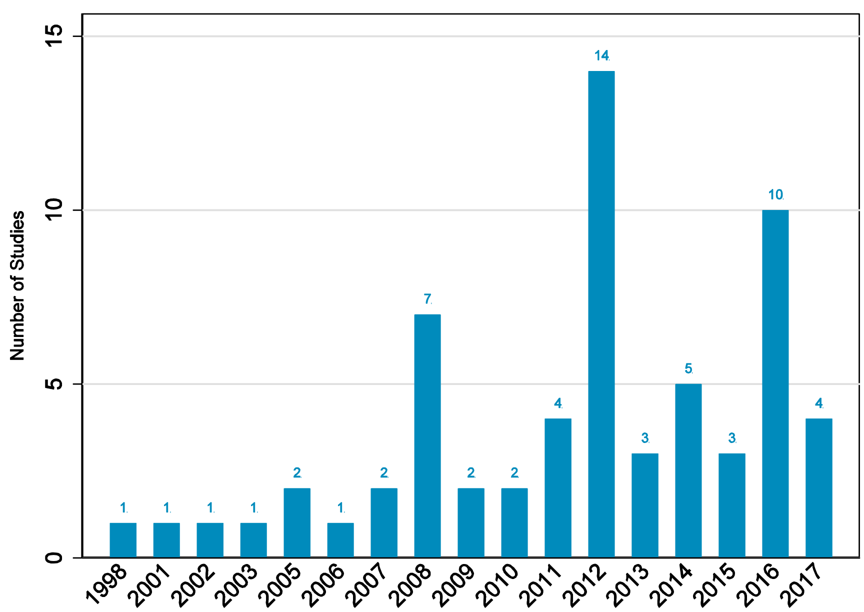
4.1. General Characteristics

With an average of 8743 words on 17 pages, the studies on the topic of HE efficiency blend into the general format of economic studies. Suitably, the majority of investigations are published in economic journals. The outlets that focus on the education sector, such as “Studies in Higher Education”, “Research in Higher Education” and “Economics of Education Review”, each having

[3] As stated in the section regarding the criteria for exclusion, we only take into account studies which assess efficiency at the institutional level. The studies by Laureti (2008) evaluating students, by Mutz, Bornmann and Daniel (2017) evaluating projects, by Cardoso and Ravishankar (2015) evaluating regions, by Titus (2009) and Titus and Pusser (2011) evaluating states, and by Hu, Yang, and Chen (2014) evaluating countries were excluded. The study of Bayraktar, Tatoglu and Zaim (2013) had to be omitted due to missing information regarding the evaluated timeframe.

published at least three studies, are most common. Nevertheless, journals which focus on empirical applications, such as “Applied Economics”, or productivity, such as the “Journal of Productivity Analysis”, are also likely to publish evaluations of efficiency of the HE sector. Examining the respective metrics of the journals, the study by Chapple et al. (2005) published in “Research Policy”, possesses the highest Impact Factor (published by Clarivate Analytics) and SCImago Journal Rank. Along with the impact of the journal, the accessibility plays an important role in the assessment of papers. To facilitate the access, the majority of studies (86%) provide keywords depicting the content. While the most frequently given attributes (“stochastic”, “efficiency” and “education”) are not surprising, the sheer number of different terms (138 attributes) is unexpected. This high value is an indication that the studies strongly differ in their aims and scenarios. Examining the publication date next, one can observe a slight increase in the number of publications over time, as illustrated in Figure 1, with the first study being published in 1998. A noticeable outlier is the year 2012, with 14 issued studies. Thereof, 12 are published by Thomas Sav, who evaluated varying characteristics of American universities relying on a detailed dataset. Accordingly, with 15 publications overall, he is the most productive author in the sample, followed by Geraint Johnes with 10, Tommaso Agasisti with six and Jill Johnes with four publications (see Table A1). Given that only two studies are published by non-natives[4], the evaluated countries are usually connected to an author of the same country. Following the argumentation of Rusnák,

Figure 1: Studies per Year

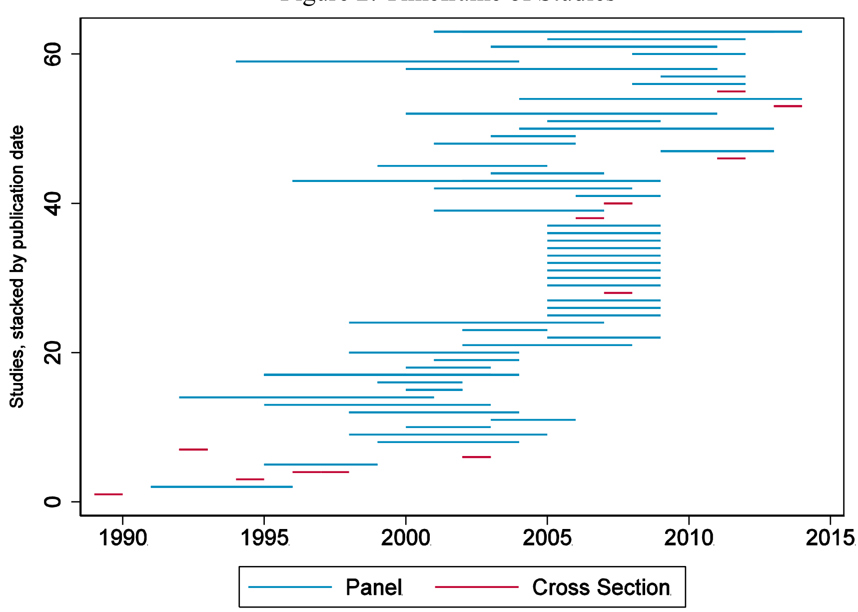


[4] A study is classified as written by a native if the stated job position of at least one author of the publication is located in the evaluated country.

Havranek and Horváth (2013), this contributes positively to the quality of the analysis, since authors are more familiar with the data at hand and are more interested in the results.

Looking more closely at the exploited data of the studies, one can observe that most studies use a panel dataset (83%, instead of a cross-section) with annual observations. The data is usually provided by the statistical office of the respective country (87%, instead of survey data or other data sources). While an average study evaluates efficiency over a five-year period, the utilized timeframes vary greatly between studies, ranging from one to 13 years. Three authors in the set additionally use partial samples to test the robustness of their results for different timeframes (Abbott and Doucouliagos, 2009; Johnes, 2014; Gralka, 2018). Figure 2 shows that one can see a slight increase in the utilized timeframe in recent years, with the cross-section remaining popular over time. Most studies choose to evaluate only public institutions (75%, instead of or in addition to private ones) and look foremost at universities (68%, instead of or in addition to colleges). This choice leads to an average evaluation of 160 institutions. The relatively high number is thereby, in particular, driven by studies on American universities and by two studies evaluating faculties (Miranda, Gramani and Andrade, 2012; Laureti, Secondi and Biggeri, 2014) instead of universities. A strong variation between studies is observable, with a minimum of six and a maximum of 954 institutions. Part of this variation is inevitably driven by the size of the HE sector of the respective countries, with 23 different nations being evaluated. While some countries, such as the United States [US], the United Kingdom [UK], Italy, Germany, Australia and Taiwan, are evaluated at least three times, the sample consists of five countries that are evaluated twice and twelve that are only considered once (see Table A1). Surprisingly, only two studies of the sample evaluate

Figure 2: Timeframe of Studies



efficiency across countries (Agasisti and Haelermans, 2016; Bolli et al., 2016).[5] Even though two further studies evaluate the efficiency of more than one country, they do so in separate estimations. The respective efficiency values in the studies by Lenton (2008) for institutions in the UK and the US and by Abbott and Doucouliagos (2009) for Australia and New Zealand are, therefore, not suited for comparison. In light of emerging comparable national datasets, e.g., the EUMIDA (2009) Dataset[6], and an increasing number of studies looking at cross-country efficiency using DEA (Rhaiem, 2017), this is unexpected. Hence, one would anticipate evaluating efficiency differences across countries, keeping in mind their individual educational systems, to be a path for future research. Interestingly, 13 studies perform an additional DEA, often to verify the resulting efficiency values.

The review of the general characteristics shows that most of the analyses blend into the format of economic studies and are published by journals focusing on the education sector. They are mostly written by natives, using panel data provided by the respective statistical office to evaluate public universities. The utilized timeframe, as well as the number of evaluated institutions, varies strongly between studies. Surprisingly the number of studies that compare efficiency across countries is low.

4.2 Underlying Assumptions

Before a specification of the SFA is chosen, assumptions concerning the type and form of the underlying function and the distribution of the efficiency term have to be made. In the upcoming section, we focus on these choices and show that there seems to be a standard in the literature.

Type of Function

In general, an SFA can be based on either a production, cost, profit or distance function.[7] The first function, i.e., the mathematical representation of the technology that transforms inputs into outputs, is thereby the most frequently used approach in most research fields. Using this type of function, researchers are able to compare one output to multiple inputs. With merely a third of all publications in our sample (32%) employing a production function, the literature on the HE sector visibly differs from other classically evaluated sectors, such as agriculture and banking. Additionally, the production function was introduced to the HE sector relatively late, with the first study using it published in 2005 (Chapple et al., 2005), the second in 2007 (Castano and Cabanda,

[5] The studies by Agasisti and Gralka (2017) and Canton, Thum-Thysen and Voigt (2018) also evaluate the efficiency across countries, using a SFA. However, both are published as working papers and are, therefore, not part of the present sample.

[6] The EUMIDA Dataset was used in a cross-country efficiency evaluation using SFA by Daghbashyan, Deiaco and McKelvey (2014). Since it is published as a chapter of a book, it is not taken into account in the above survey.

[7] It is generally assumed that the quality of all outputs and inputs are similar, universities use the same technology and there is linear homogeneity with respect to input factor prices.

2007) and its more repeated use only starting in 2012. While the production function allows to assess efficiency in a technical sense, evaluating whether the inputs are fully utilized given the technology, it does not permit assessing whether the observed combination of inputs is the best. Assuming that a unit can produce the same level of output using different input combinations, a criterion is needed to judge which combination is most advisable. Therefore, technical efficiency can be complemented by allocative efficiency, which represents the appropriate ratios of inputs. The extension is possible when a cost function is used as the underlying function.[8] The cost function can be seen as a boundary describing the lowest cost at which an institution can produce a set of outputs. It entails the comparison of expenditures to outputs and input prices.[9] Following most authors, among them Eagan and Titus (2016), our survey shows that a cost function is customarily used to estimate efficiency in the HE sector. This makes sense, because universities are multi-output institutions, typically engaging in teaching and research. Hence, a production function that examines only one aspect of university production can be problematic. Interestingly, the cost function is not only the most frequently used function (41%) but also the one which was used almost solely in earlier publications. Apart from the production and cost function, distance functions have gained popularity in the last years. This development is driven by their advantage that multiple inputs and outputs can be included.[10] In the publication timespan from 2014 to 2017, ten studies (27% of the sample) were published using these types of functions for the HE sector (Abbott and Doucouliagos, 2009; Johnes, 2014; Nemoto and Furumatsu, 2014; Olivares and Wetzel, 2014; Erkoc, 2015; Kulshreshtha and Nayak, 2015; Agasisti, Barra and Zotti, 2016; Bolli et al., 2016; Titus, Vamosiu and McClure, 2016; Barra, Lagravinese and Zotti, 2018). Most likely, due to the unavoidable mismatch concerning the number of input and output variables, there is no study evaluating the effects of different types of functions on the resulting efficiency values for the HE Sector. [11]

[8] Alternatively, allocative efficiency can be introduced through a profit function, assuming profit maximization as the underlying behavioral criteria. Since the form is not applied in an evaluation of the HE sector, we refrain from discussing it further.

[9] Notably, there are cases where either input prices do not vary much (consistent with a competitive market) or the input price data is not available. In such cases, a cost function can be estimated without the inclusion of prices, concentrating on technical efficiency only.

[10] For simplicity, multi-product cost functions, such as the one used by Titus, Vamosiu and McClure (2016), are classified as a distance function in the present evaluation.

[11] Recent literature on SFA evaluations, among them Kumbhakar, Wang and Horncastle (2015), emphasize the necessity to test the data with an OLS production (cost) regression model beforehand. If the estimated values of the random error are negatively (positively) skewed, technical inefficiency may be evident in the data. A majority of studies (64%) follow the advice and discuss the skewness of the OLS residuals.

Functional Form

Along with the type of function, its functional form plays an important role in the estimation of efficiency. The literature emphasized the difficulty of choosing an appropriate functional form and highlighted four that make sense in the general multiproduct context, i.e., the linear, the quadratic, the constant elasticity of substitution and the translog specification (Baumol et al., 1982). The first, i.e., the linear functional form, is the simplest, although it does not allow to consider and, therefore, evaluate interactions between factors. Hence, it is used in merely 4% of all estimations of efficiency in the HE sector. The second, the quadratic functional form, has the relative advantage that it is well defined for zero values. However, it is only occasionally employed in efficiency evaluations, mainly because it is not possible to ensure linear homogeneity in prices without a further normalization. In the literature of the HE sector, only 14% of all estimations use the quadratic form. The third, the constant elasticity of substitution and, in particular, the Cobb-Douglas functional form, is more frequently used, due to its straightforward equation and requirement to estimate only a few parameters. Nevertheless, the form is known to present some conceptual difficulties and authors need to make further assumptions regarding the elasticities of substitution (Johnes, 2004; Titus, Vamosiu and McClure, 2016). Among the evaluated studies, only one study chose to use the constant elasticity of substitution, and 32% of all estimations used the Cobb-Douglas functional form. The fourth, the translog function, is demanding both in terms of data and its highly non-linear specification. However, it has the advantage of having a sufficiently flexible form and no assumptions regarding the elasticities of substitutions have to be made. A further benefit of the specification is the valuable information offered by the included cross-terms, although the parameters cannot be directly interpreted due to their non-linear nature (Coelli et al., 2005). These advantages are reflected in the number of applications, with 45% of all estimations choosing the translog function. Unexpectedly, for 4% of all estimations, the respective authors did not specify the underlying functional form used for the estimation.

Observing these variations concerning the functional form in the literature on HE efficiency, Eagan and Titus (2016) evaluated their impact on the resulting efficiency values. Using the HE sector of the US as an example, they show that the chosen form slightly influences the mean as well as the distribution of the efficiency values. These results hold true for panel as well as cross-section data, for varying distributional assumptions and across different SFA specification. The authors therefore suggest to always test various forms when evaluating the efficiency of universities. A recommendation that five studies already fulfilled or followed later, most often to show the robustness of their results. While Agasisti and Belfield (2017) contrast the linear and Cobb-Douglas functional form, Chapple et al. (2005), Sav (2011), Agasisti, Barra and Zotti (2016) and Barra, Lagravinese and Zotti (2018) compare efficiency values for the Cobb-Douglas and the translog form. In all five cases, the resulting efficiency values show a high similarity.

Distribution of the efficiency term

The third underlying assumption concerns the distribution of the efficiency term. Thereby, the opinion regarding its effect on the efficiency values varies greatly. While Eagan and Titus (2016) state that the choice is "critical to a stochastic frontier analysis" (Eagan and Titus, 2016, p. 448), Coelli et al. (2005) point out that the rankings based on efficiency estimates are often quite robust to the distributional assumption, even if the values themselves differ. In principal, the efficiency term could follow any non-normal distribution, so that it can be separated out from the noise term. Common assumptions are that it follows a half-normal, exponential or truncated normal distribution.[12] The first two are thereby recommended most often (Coelli et al., 2005). They have a modal value at zero, implying that inefficiency is close to zero and the associated values of technical efficiency are close to one. Both are used in studies on the HE sector, with 40% of all estimations choosing a half-normal and 5% choosing an exponential distribution. Alternatively, a truncated normal distribution can be assumed, which allows for a wider range of distribution shapes. Unfortunately, this sort of flexibility comes at the cost of computational complexity. Therefore, slightly fewer estimations in our sample, approximately 32%, use the truncated distribution. Notably, the choice regarding the distributional assumptions is discussed surprisingly little in the evaluated papers, and for 19% of all estimations in our sample, no information regarding the assumed distribution of the efficiency term is given.

In their comprehensive evaluation of SFA variations for the HE sector, Eagan and Titus (2016) show that the assumption on the error distribution has almost no impact on the efficiency values. This result is verified in four further studies, with Sav (2012l), Johnes (2014) and Agasisti and Belfield (2017) showing the robustness of their results for the truncated and half-normal and by Erkoc (2015) for the half-normal and exponential distribution. The studies by Horne and Hu (2008), Johnes (2014) and Laureti, Secondi and Biggeri (2014) are notable exceptions and even avoid the specific parametric assumptions concerning the error terms (see section 4.3).

Summarizing, there seem to be standard assumptions when estimating efficiency of HE institutions, with a majority of studies choosing to evaluate a translog cost-function, assuming a half-normal distribution of the efficiency term. Regrettably, authors often miss the opportunity to test and explain their choices regarding the type and form of the function and the distributional assumptions.

---

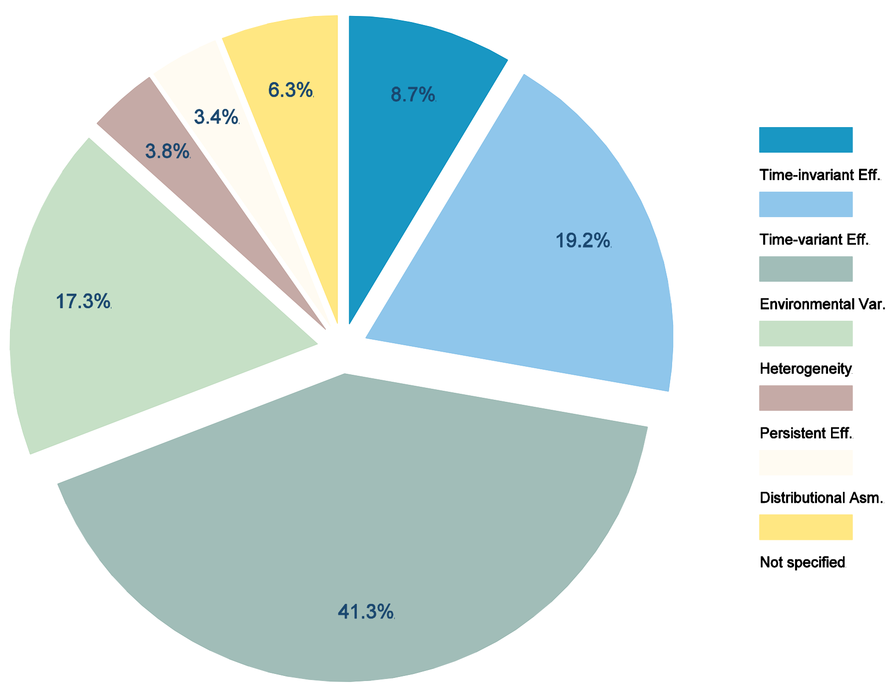
[12] Common in the efficiency literature is also the gamma distribution. Since it is not applied in the HE literature, we do not discuss it further.

4.3 Applied Specifications

A wide variety of SFA specifications exist. The objective of the following description is to give an overview of specifications that were applied to the HE sector, explaining their relative advantages and shortcomings. In each case, the specification is briefly described, before studies employing it are presented. Depending on the frequency of their application, each approach is explained in more or less detail. Naturally, the chapter does not replace an in-depth study of the mentioned papers and methods. Corresponding to their methodical focus, the specifications can be grouped into the following six classes: approaches that assess time-invariant efficiency, time-variant efficiency, environmental variables, heterogeneity, persistent inefficiency or distributional assumptions. Figure 3 displays the share of these classes in estimations of HE institutions. Evidently, specifications that evaluate environmental variables, time-variant efficiency and heterogeneity are most frequently employed. Hence, the graph reflects that there seems to be somewhat of a standard in the literature regarding the efficiency of HE institutions. This is confirmed in the following review, with two specifications, i.e., environmental variables and time-variant efficiency, used most frequently. Nevertheless, the most prominent development concerns the consideration of heterogeneity, as there are a large variety of applied specifications. Regrettably, for 13 estimations in our sample, the employed specification was not specified.

To make the following overview as transparent as possible, we introduce a baseline model, which consists of four equation lines. If a whole class (or specification) differs from the baseline model,

Figure 3: Classes of Specifications (Count of Estimations)



the respective line is shown in its adjusted form. Since the majority of studies on the HE sector apply a cost function, we chose to use the functional form as an example henceforth. Based on Aigner, Lovell and Schmidt (1977) the panel data version can be written in the following general form:[13]

$$C_{it} = f(y_{it}; \beta) + \varepsilon_{it} \qquad (1.0)$$

$$\text{with} \quad \varepsilon_{it} = v_{it} + u_{it} \qquad (2.0)$$

$$u_{it} \geq 0 \qquad (3.0)$$

$$v_{it} \sim N(0,\, \sigma_v^2) \qquad (4.0)$$

$$i=1,\dots,N;\ t=1,\dots,T$$

The equations show a general cost function, where C_{it} represents the costs of university i at time period t. The function $f(\mathrm{y}_{it}; \beta)$ describes the output technology. The composed error term ε_{it} consists of v_{it} and u_{it}, which are both distributed independently of each other and of the regressors. The former accounts for the normally distributed noise term. The latter represents the non-negative random error term, which captures technical inefficiency.[14] The parameters of the model are estimated in two sequential steps. In the first step, the estimates of the model parameters are obtained by maximizing a log-likelihood function [ML]. Since the estimates of the model parameters allow the computation of the error term ε_{it}, but not the efficiency estimates, a second step is necessary. Exploiting the conditional distribution of u_{it} given ε_{it}, one can either calculate inefficiency, based on the approach by Jondrow et al. (1982) [JLMS] or efficiency, relying on Battese and Coelli (1988) [BC].

Time-Invariant Efficiency

The earliest specification which was applied in an estimation of HE efficiency is the random effects specification by Pitt and Lee (1981). The authors assume that efficiency is constant over time and follows a half-normal distribution, shown in equations (2.1) and (3.1a). While the specification was the first to extend existing approaches to include longitudinal data, it has two main disadvantages. First, as Kumbhakar and Lovell (2003) point out, the underlying timeframe needs to be sufficiently large to obtain consistent estimates of the efficiency term. However, if this is the case, the assumption that efficiency does not change over time is implausible. Second, it is assumed that there are no structural differences (heterogeneity) between the firms. Farsi, Filippini and Greene

---

[13] The numbering of the equations corresponds to the equation line (first figure) and specification (second figure).

[14] For simplicity, a case with time-variant inefficiency is shown as a baseline.

(2006), among others, show that it can, therefore, be seen as a lower boundary of efficiency, where all individual, time-invariant effects are categorized as inefficiency. Modifying the baseline model, the Pitt and Lee (1981) specification can be written as follows:

$$with \quad \varepsilon_{it} = v_{it} + u_i \qquad (2.1)$$

$$u_i \sim N^+(0, \sigma_u^2) \qquad (3.1a)$$

The specification by Pitt and Lee (1981) is used in 12 estimations within 6 studies in our sample, ranging from the publication dates of 2002 to 2016 (Izadi et al., 2002; Lenton, 2008; Das and Das, 2014; Erkoc, 2015; Agasisti and Haelermans, 2016; Zhang, Bao and Sun, 2016).

While Pitt and Lee (1981) suggested a half normal distribution of the error term, Battese and Coelli (1988) proposed a generalization of this model, assuming a truncated normal distribution, shown in equation (3.1b):

$$with \quad u_i \sim N^+(\mu, {\sigma_u}^2) \qquad (3.1b)$$

This second specification to measure time-invariant efficiency by Battese and Coelli (1988) is used in 6 estimations within 4 of the studies of our sample (Fu, Huang and Tien, 2008; Johnes, 2014; Nemoto and Furumatsu, 2014; Zhang, Bao and Sun, 2016).

Time-Variant Efficiency

Since the assumption that technical efficiency is constant through time is rather restrictive, a variety of approaches emerged, allowing efficiency to change over time. Thereby, the simplest solution is probably to treat each year as a cross-section, forgoing the advantages of panel data. This so called "pooled model" was applied for the HE sector by Laureti, Secondi and Biggeri (2014), Erkoc (2015) and Zhang, Bao and Sun (2016). However, in all three studies, the method was just seen as a baseline, which allowed the verification of more advanced approaches. In addition to the pooled model, specifications were also proposed that allowed the estimation of time-varying efficiency. The majority of studies, thereby have a common framework, shown in equations (3.2). It is assumed that inefficiency u_{it} is composed of the following two distinct components: a stochastic individual component u_i, which is constant over time, and a non-stochastic function of time G_t that is common for all institutions. While these specifications allow efficiency to vary over time, the change is subject to a fixed time pattern, which is assumed to be the same for all institutions. Again, heterogeneity is not considered within the model. The baseline model can be adjusted to:

$$\begin{aligned} with \quad & u_{it} = G_t * u_i \, , G_t \geq 0 \\ & u_i \sim N^+(\mu, {\sigma_u}^2) \end{aligned} \qquad (3.2)$$

In the evaluation of HE efficiency two time-varying specifications are applied. The first is the specification by Kumbhakar (1990), who proposes a function of time G_t which includes two parameters, shown in equation (3.2a). The specification is only used once, by Zhang, Bao and Sun (2016), with the aim to verify the efficiency values from a different specification. Adjusting the baseline model, the Kumbhakar (1990) specification is as follows:

$$G_t = [1 + exp(\gamma_1 t + \gamma_2 t^2)]^{-1} \qquad (3.2a)$$

In contrast, the similar, subsequently proposed "Time Decay Model", by Battese and Coelli (1992) is more frequently used. It differs from the former only in the specific form of the time-varying component. Including only one parameter and T, the terminal period of the sample, the function of time proposed by Battese and Coelli (1992) is shown in equation (3.2b):

$$G_t = exp[\gamma(t - T)] \qquad (3.2b)$$

The specification is used in 35 estimations within 18 studies in our sample and is, therefore, the second most frequently used approach for the evaluation of the HE sector. While some authors use the specification as the main approach (Sav, 2012b; Kulshreshtha and Nayak, 2015; Agasisti, 2016), the majority of authors employ it as a baseline to verify other (often more advanced) specifications (McMillan and Chan, 2006; Johnes and Johnes, 2009; Agasisti and Johnes, 2010; Johnes and Schwarzenberger, 2011; Sav, 2011; Sav, 2012c; Sav, 2012d; Sav, 2012e; Johnes, 2014; Laureti, Secondi and Biggeri, 2014; Agasisti and Johnes, 2015; Erkoc, 2015; Zhang, Bao and Sun, 2016; Agasisti and Belfield, 2017; Gralka, 2018).

Studies comparing results from the aforementioned time-invariant and time-variant specifications, namely, Johnes (2014), Erkoc (2015) and Zhang, Bao and Sun (2016), conclude that both lead to an overall similar assessment of institutions.

Environmental Variables

In addition to the level of efficiency, researchers are interested in the factors that can explain inefficiency. Hence, specifications emerged that allow the incorporation of exogenous variables, i.e., the so called "z-variables" or "determinants of efficiency". These are neither inputs nor outputs of the production process and are outside of the control of the institutions but are assumed to influence the producer performance nonetheless (Coelli et al., 2005). The employed determinants are discussed in chapter 4.5. The first advancement to investigate the relationship between efficiency and its determinants was a two-step procedure. The individual efficiency values, which are estimated in the first step, are thereby regressed on a vector of exogenous variables in a second step. Depending on the resulting coefficient of the regression, the determinant either increases or decreases efficiency. However, the two-step procedure was soon recognized as biased. If the inputs or outputs from the efficiency evaluation are correlated to the determinants of efficiency, the ML estimates and, therefore, the efficiency values are biased, due to the omission of relevant variables (Wang and Schmidt, 2002). Additionally, Battese and Coelli (1995) point out that the assumption from the first step, i.e., that inefficiencies are identically distributed, is contradicted in the second step when a regression model is specified for the predicted efficiency values. Fortunately, the specification is used in only four estimations within 2 studies in our sample (Mensah and Werner, 2003; McMillan and Chan, 2006).

Given the criticism, specifications were proposed to study exogenous determinants of efficiency in a single step procedure. In this case, z-variables are estimated together with all the other parameters in the ML estimation. While there are numerous specifications published (see Kumbhakar, Wang and Horncastle (2015) for a compact overview), the version by Battese and Coelli (1995) is the most prominent one.[15] The authors abandon the assumption of an efficiency term with a constant mean. Instead, they assume that the mean is a linear function of the exogenous variables, as shown in equation (3.3). Thereby z_i is a vector of the exogenous variables of observation i and δ is the corresponding coefficient vector. The specification not only allows the inclusion of exogenous variables but also makes the distributional assumption of u_i more flexible, since each observation has an individual mean. Adjusting the baseline model, the Battese and Coelli (1995) specification can be written as follows:

$$\begin{aligned} &with \quad u_{it} \sim N^{+}(\mu_{it}, {\sigma_u}^2) \\ &\mu_{it} = z_{it}'\delta \end{aligned} \qquad (3.3)$$

The specification is used in 78 estimations within 25 studies in our sample and is, therefore, the most frequently used approach in the HE Sector. The majority of authors use it as the main specification and focus on the selection and interpretation of the chosen z-variables, sometimes performing multiple estimations, with varying variables and determinants (Robst, 2001; Chapple et al., 2005; Stevens, 2005; Castano and Cabanda, 2007; Kuo and Ho, 2008; Johnes, Johnes and Thanassoulis, 2008; Kempkes and Pohl, 2008; Abbott and Doucouliagos, 2009; Kempkes and Pohl, 2010; Mamun, 2011; Sav, 2012d; Sav, 2012f; Sav, 2012g; Sav, 2012h; Sav, 2012i; Sav, 2012j; Sav, 2012k; Sav, 2012l; Sav, 2013; Bolli et al., 2016; Sav, 2016). However, some studies additionally contrast the results from the specification to other approaches (Sav, 2011; Sav, 2012c; Sav, 2012e; Laureti, Secondi and Biggeri, 2014; Agasisti and Belfield, 2017), which are discussed in the following segments. A similar approach by Kumbhakar, Ghosh, and McGuckin (1991),

[15] Some of the subsequently proposed specifications adopt the approach by Battese and Coelli (1995) and likewise assume that the mean is a linear function of exogenous variables. This includes the specifications by Wang and Ho (2010) as well as the true random and true fixed effects specifications.

which is, however, limited to cross-sectional data, was used by McMillan and Chan (2006) to assess the efficiency of Canadian institutions.

Another specification that also focuses on the inclusion of determinants of inefficiency was proposed by Laureti (2008). It challenges the assumption that both error terms are homoscedastic, which implies that the variance of the noise and inefficiency parameters are constants, and instead, proposes a heteroskedastic inefficiency term. The variance of the inefficiency term is a function of exogenous variables and, therefore, is allowed to vary across institutions. Compared to the specification by Battese and Coelli (1995), it is, therefore, not the mean, but the variance of the efficiency term that depends on the “z-variables”. The specification is used solely in the study by Zoghbi, Rocha and Mattos (2013) evaluating the HE sector of Brazil. The authors perform six estimations in total, varying the evaluated sample and the considered environmental variables.

Heterogeneity

A major concern of researchers when measuring the efficiency of universities, concerns the structural differences that undoubtedly exist between them. Institutions feature different locations, fields and extents of research as well as governance structures and, therefore, a high degree of heterogeneity exists. The literature thereby defines the term heterogeneity as permanent differences between institutions, which cannot be altered by them and should, therefore, be ruled out from the efficiency term. A lack of consideration can lead to biased results, since the error and, hence, the efficiency term might pick up these structural differences. Therefore, specifications that do not account for heterogeneity assumedly underestimate efficiency. To account for heterogeneity in the HE sector, the literature initially focused on homogenous groups of institutions and evaluated only universities, leaving out polytechnics as well as all specialized and private institutions. Alternatively, some authors, such as Johnes et al. (2005), estimated cost functions specific to certain prespecified subgroups of institutions. However, these procedures have drawbacks. The a priori classification requires detailed knowledge on each institution and might nevertheless be unsuitable. Additionally, as Kumbhakar, Wang and Horncastle (2015) point out, the separation could lead to biased estimations of efficiency. Institutions of different groups still share common features that could be relevant for the accurate estimation of the frontier. To account for those shortcomings, specifications that allow heterogeneity to be included directly in the estimation were applied for the evaluation of the tertiary sector, soon after their general introduction.[16] Those include the metafrontier approach, the latent class estimation, the true effects models, the specification by Wang and Ho (2010) and the random parameter model, all presented in the following section.

[16] An extended review of these models can be found in the survey by Greene (2008).

One of the first specifications to account for heterogeneity is the so called "metafrontier approach", proposed by Battese, Rao, and O'Donnell (2004). In a first step, the overall sample is divided into subclasses, grouping institutions by their technology. The efficiency is then estimated separately for each group. Afterwards, the respective production frontiers of each group are compared. The best practice frontier is seen as the meta frontier and the differences across groups are viewed as technology gaps. The approach, therefore, not only allows the efficiency of an institution relative to the technology it uses to be estimated but also the evaluation of the technology gap relative to the best practice technology. While the specification still requires an a priori classification, it allows different technologies to at least be compared. The metafrontier approach is used in only one study on the HE sector. Lu and Chen (2013) compare the efficiency of Taiwanese Institutes and Universities of Technology using a metafrontier cost-function framework.

Alternatively, authors deeming the categorization of institutions reasonable but not straightforward, rely on the latent class estimation. Latent classes are thereby defined as unobservable (~latent) subgroups (~classes) of units that are homogeneous in certain criteria. While different latent class stochastic frontier models (LCSFM) exist, they have a three-step approach in common.[17] First, following the general idea of latent class models, each institution is assigned multiple weights, which reflect the probability of a membership in each class and are obtained through a likelihood function. In the second step, as many frontiers as number of classes are estimated. Third, the individual efficiency is estimated either by assigning each institution a class based on the highest probability (ignoring all other class probabilities) or the efficiency value is weighted according to all class probabilities (Orea and Kumbhakar, 2004). Nevertheless, since frontier functions are estimated separately for each class within the LCSFM framework, not all potentially relevant information is used for the generation of the frontiers. Additionally, the number of classes has to be predefined by the analyst. The specification is used in two studies in our sample. In an evaluation of the English HE sector, Johnes and Johnes (2016) apply a LCSFM with two classes. The class allocation shows that the size of institutions plays an important role, with one group comprising larger, research-intensive universities and one containing smaller, more specialist institutions. While they contrast the parameter estimates of each output from the LCSFM with those of a "straightforward SFA" (Johnes and Johnes, 2016, p. 605), they neither give details regarding the latter specification nor discuss differences or similarities in depth. In contrast, Agasisti and Johnes (2015) also predefine two classes in their evaluation of US institutions, but they provide an in-depth comparison of the efficiency values to those of a standard time-variant

[17] Well-known models are put forward by Caudill (2003) and Orea and Kumbhakar (2004). A detailed overview can be found in Greene (2005b). The three steps are not necessarily successively implemented; for example, Orea and Kumbhakar (2004) propose a single stage estimation.

specification. As they expect, the LCSFM classifies more institutions as efficient, since colleges are evaluated only to similar competitors. With the classes main difference being the size of institutions again, the authors argue that the specification allows the different scales of operation to be considered.

A different approach to include heterogeneity directly into the efficiency estimation was put forward by Greene (2005a) in the form of the "true effects models". He proposes to include a university-specific, time-invariant component ρ_i in the time-varying model, see equations (1.4) and (3.4). In doing so, it is assumed that all constant influences revert to heterogeneity. Since the model allows inefficiency to be separated from institution specific heterogeneity, the specification is known as the "true fixed effects" or the "true random effects" model, depending on the treatment of the time-invariant term as fixed or random.[18] While the estimation of the true random effects specification can be performed using simulation-based techniques, the estimation of the fixed effects specification brings along some difficulties, the most prominent being the incidental parameter problem (Neyman and Scott, 1948). It arises when the number of estimated parameters increases with the number of cross-sections in the data. Greene (2005a) shows that the solutions for the standard panel data model, as the conditional or marginal ML, are not applicable for stochastic frontier models. He proposed a provisional solution by including N dummy variables in the model. In both versions of the specification (fixed and random), it is possible to consider z-variables. Although the specification provides a straightforward way to account for heterogeneity, it can be argued that not all individual, time-invariant effects should be classified as such (see the following discussion on persistent inefficiency). Accordingly, Farsi, Filippini and Greene (2006), among others, show that the results can be seen as an upper boundary of efficiency. The Greene (2005a) specification can be written as:

$$C_{it} = f(x_{it}; \beta) + \rho_i + \varepsilon_{it} \tag{1.4}$$

$$\begin{aligned} \textit{with} \quad & u_{it} \sim N^+(0, \sigma_u{}^2) \\ & \rho_i \sim N(0, \sigma_\rho{}^2) \end{aligned} \tag{3.4}$$

Within our sample, Fieger, Villano and Cooksey (2016) use the true fixed effects specification and Olivares and Wetzel (2014) as well as Bachan (2017) use the true random effects version as the main evaluation method. In contrast, Gralka (2018) compares the true random and Laureti, Secondi and Biggeri (2014) and Zhang, Bao and Sun (2016) compare the true random and true fixed specifications to other approaches. The three studies confirm the results from Farsi, Filippini and

[18] A similar inclusion of a time-invariant term is proposed by Kumbhakar and Heshmati (1995), who interpret the term as persistent inefficiency instead of heterogeneity. Since the specification was not applied to the HE sector, it is not discussed further.

Greene (2006) for the HE sector, showing that the true effect models lead to higher efficiency values than the specifications by Pitt and Lee (1981), Battese and Coelli (1988), Kumbhakar (1990), Battese and Coelli (1992), Battese and Coelli (1995) and the pooled model.

Wang and Ho (2010) later proposed a more consistent solution for the true fixed effects specification. They propose an SFA specification, which builds upon a model by Wang and Schmidt (2002). Therefore, within or first-difference transformations are used to avoid the incidental parameters problem. The model transformation is possible due to the multiplicative form of inefficiency, combining individual-specific effects u_i and individual and time-specific effects h_{it}. The latter is thereby defined as a positive function of the non-stochastic inefficiency determinants z_{it}. Employing the cost function (1.4) proposed by Greene (2005a), the efficiency term is assumed to follow a truncated normal distribution. The specification is shown in equation (3.5):

$$\begin{aligned} \textit{with} \quad u_{it} &= h_{it} * u_i^* \\ h_{it} &= f(z_{it}\delta) \\ u_i^* &\sim N^+(\mu, {\sigma_u}^2) \end{aligned} \tag{3.5}$$

The specification is used in 15 estimations within two studies in our sample. Agasisti, Barra and Zotti (2016) as well as Barra, Lagravinese and Zotti (2018) use the approach to assess the Italian HE sector. They evaluate the robustness of their results by varying the functional form of their distance function, by including interaction terms between the input and output variables and by employing diverse determinants of efficiency. Since neither of the studies compare the results of the specification to other approaches, a statement regarding the comparability of the results of the Wang and Ho (2010) specification to other approaches is not possible.

Together with the true effects models, Greene (2005a) proposed a further alternative approach to include heterogeneity. The so called "random parameters model" relaxes the assumption that all units must face the same underlying cost function. Instead, the specification outlined in (1.6) and (3.6) allows separate functions for each institution to be estimated. In comparison to the true effects model, where only the constant ρ_i is allowed to vary across institutions, the random parameter model allows the other parameters of the function, namely, α_i, β_i and μ_i and θ_i, to vary as well. However, they are constrained to follow a specified statistical distribution. A simulated maximum likelihood is used to estimate the parameters of the model. The random parameters model specification can be written as:

$$C_{it} = \alpha_i + f(x_{it}; \beta_i) + \rho_i + \varepsilon_{it} \tag{1.6}$$

$$\text{with} \quad u_{it} \sim N^{+}(\mu_i, \sigma_{ui}^2)$$
$$\mu_i = \mu_i' z_i \qquad (3.6)$$
$$\sigma_{ui} = \sigma_u exp(\theta_i' h_i)$$

The specification is used in six studies in our sample, with one estimation performed in each. While it is employed as the sole approach in the studies by Johnes and Salas-Velasco (2007) and Johnes, Camanho and Portela (2008), the majority of studies additionally contrast the results to other specifications. In an evaluation of the Italian HE sector, Agasisti and Johnes (2010) compare it to the Pitt and Lee (1981) specification and show that the efficiency values from the earlier specification are more dispersed than the ones from the random parameter specification. Nevertheless, the correlation between the values is quite high, as shown by a high spearman rank correlation. While Johnes and Johnes (2009) and Johnes and Schwarzenberger (2011) also use the time-invariant approach as a baseline, they only show the estimated parameters of both specifications and do not discuss similarities or differences further. The study by Agasisti and Johnes (2015), who additionally employ the LCSFM, is informative. They show that the new specification leads to higher efficiency values than the classic approach and the LCSFM. They argue that this is mainly driven by the fact that with each specification, the respective competitors become more similar; in the classic specification, all institutions are compared to one frontier, while in the latent class approach, institutions are grouped and two frontiers are evaluated, and in the random parameters model, each institution is compared with their own (potential) performance.

Heterogeneity and Persistent Inefficiency

Concurrent with the proposal to include heterogeneity in the evaluation, a specification was proposed that aimed to separate short-term (~transient) and long-term (~persistent) efficiency (Kumbhakar and Heshmati, 1995). The short-term efficiency thereby reflects changes that occur in a given year, while the long-term efficiency echoes the effects of constant influences. While it was not applied to the HE sector, a more recent specification that combines both arguments is.[19] It is thereby assumed that it is misleading to assign all individual, time-invariant effects to unobserved heterogeneity. Instead, a fraction of those effects belongs to inefficiency. The management, for example, differs between the institutions and is commonly a long-term factor. However, it should not be seen as heterogeneity, as it is, in fact, adaptable in the long run. Therefore, it is insufficient to include only one individual, time-invariant term in the estimation equation. It is necessary to distinguish between heterogeneity and long-term inefficiency. Consequently, a four components

[19] The specification by Kumbhakar and Heshmati (1995) was applied to the Italian and German HE sector by Agasisti and Gralka (2017). Since the analysis is only available as a working paper, it is not included in the present survey.

specification is proposed that allows transient and persistent efficiency to be estimated while controlling for heterogeneity and noise. The idea of the specifications is outlined in equations (1.7) and (3.7). Thus, the term ρ_i is again a random institution effect that captures heterogeneity. The overall inefficiency term u_{it} is divided into a transient component τ_{it} and a persistent part μ_i. Four different variants of the four-component specification exist, which vary in their estimation procedure. Kumbhakar, Lien, and Hardaker (2014) proposed a multistep procedure, while Tsionas and Kumbhakar (2014) put forward a Bayesian method. In contrast, both Colombi et al. (2014) and Filippini and Greene (2016) use a single step procedure, the former using an ML and the latter a maximum simulated likelihood for the estimation. An accomplished discussion on the approach can be found in Badunenko and Kumbhakar (2016). The specification by Kumbhakar, Lien, and Hardaker (2014) is given by:

$$C_{it} = f(x_{it}; \beta_i) + \rho_i + \varepsilon_{it} \qquad (1.7)$$

$$\begin{aligned} \textit{with} \quad & u_{it} = \mu_i + \tau_{it} \\ & \mu_i \sim N^+(0, {\sigma_\mu}^2) \\ & \tau_{it} \sim N^+(0, {\sigma_\tau}^2) \\ & \rho_i \sim N(0, {\sigma_\rho}^2) \end{aligned} \qquad (3.7)$$

The specification is used within two studies in our sample. Both argue that the specification allows for a more accurate estimation of heterogeneity and, thus, efficiency. In an evaluation of US institutions, Titus, Vamosiu, and McClure (2016) use a slightly modified version of the Kumbhakar, Lien, and Hardaker (2014) specification and additionally take spatial interdependency into account. They show that cost inefficiency tends to be persistent rather than short term for master's institutions in the US. This is later confirmed for the Germany HE sector by Gralka (2018). The latter author additionally compares the results of the model to the Battese and Coelli (1992) and the true random effects specification. Compared to the former, the newer specification leads to (the expected) higher efficiency values. Compared to the latter, the specification leads to (the likewise expected) lower efficiency values. This is in line with the argument that it is necessary to account for heterogeneity, but to not classify all time-invariant effects as such. The correlation of the university rankings is significantly positive but partially low, indicating that university-specific conclusions are likely to vary by method.

Distributional Assumptions

In contrast to the afore noted publications on the efficiency of the HE sector, three studies in our sample estimate efficiency without imposing distributional assumptions on the error term. The earliest specification thereby focuses on time-invariant efficiency. Schmidt and Sickles (1984) propose an alternative approach to the usually employed ML method for the estimation, namely, the generalized least squares [GLS] technique. In our sample of studies evaluating the HE sector, it was only used once by Johnes (2014), with the aim of verifying the results from an estimation based on the Battese and Coelli (1988) specification (time-invariant efficiency) and the Battese and Coelli (1992) specification (time-variant specification). He shows that the estimation leads to slightly lower efficiency values than the two more popular ones.

Another specification that allows distributional assumptions to be avoided, in this case, for time-varying efficiency, was proposed by Cornwell, Schmidt and Sickles (1990). They do so by introducing a flexible function of time in the production function. However, Kumbhakar, Wang and Horncastle (2015) point out that in a setting where a large number of institutions are evaluated for a short timeframe, the specification is quickly overparameterized. Within our sample, it was only used once by Horne and Hu (2008) in an evaluation of the Australian HE sector. They do not contrast the results of the specifications with other estimations.

In a third study in our sample, a novel approach is introduced. With the main purpose to avoid distribution assumptions, Laureti, Secondi and Biggeri (2014) introduce the generalized maximum entropy [GME] estimator into the SFA. In their evaluation of Italian institutions, the authors compare the results from the GME to the pooled model and the time-invariant and the true fixed effects specification. They conclude that the efficiency estimates obtained by the novel approach are lower than those obtained with the specifications that are based on the ML method. Most likely, due to the complex implementation, not many studies exist that use or even evaluate the specification. In a more general assessment of the approach, Campbell, Rogers and Rezek (2008) even consider it not as a development of the SFA, but as an alternative frontier estimation approach to the known SFA and the DEA.

The review of applied specifications shows that the specifications by Battese and Coelli (1995), including determinants of efficiency, and by Battese and Coelli (1992), estimating time-variant efficiency, are the two most frequently used approaches in the evaluation of HE institutions. The most prominent development concerns the consideration of heterogeneity, with a large variety of applied specifications. A recent advancement concerns the distinction between short- and long-term efficiency, which arguably allows for a more accurate measurement of heterogeneity. An interesting, not yet exhausted topic concerns the avoidance of the usually necessary distribution assumptions.

4.4 Utilized Factors

In the following section, an overview of the inputs and outputs that were used in the efficiency evaluations of universities is given. Afterwards, the employed input prices and dummy variables are reviewed.[20] We show that, even though there seems to be a core model representing the expenditures as well as the teaching and research tasks of universities, the choice of variables complementing those is diverse. The following description focuses on the estimation level since the variables are often varied within one study.

Before proceeding with the overview, a note on the general challenges that occur when identifying the factors to measure the productivity of universities is necessary. The HE sector is multifaceted and the variables chosen can only ever approximate the real situation. A main difficulty is presented by the circular meaning of some core variables. For instance, the number of students can be considered as a teaching output but can also represent an input if the number of graduates is also included. A similar problem exists for research grants, which can be seen as an input for the research production, but within efficiency evaluations, they often represent an approximation of research output. A further challenge concerns the timing of variables. It is assumed that all factors are connected within one year. To give an example, for a cost function, this assumption implies that the expenditures of a given year are associated with the graduations and publications of the same year. However, both outputs are presumably influenced by the expenditures of the years before. In addition, the quantity of the employed variables is restricted within the SFA and depends on the type of function (see section 2). For the evaluation of employed factors, it is, therefore, necessary to classify studies according to the functional form they use. The examination of the input category "personnel structure" demonstrates this visibly. While it is employed in only 52% of all estimations, in fact 83% of all production function and 98% of all distance function estimations use it. The relatively low overall number is driven by the fact that employees are not used in any cost function, since in this type of function, only one input, the expenditures, can be included. In the following assessment, we, therefore, emphasize the employed functional form.

Input Factors

To provide a clear overview, we grouped the employed inputs into six categories relating to budget, personal structure, students, physical capital, research funds and, finally, other factors. They are presented according to the frequency of their usage. On average, six input factors are used in a production function, one in a cost function and three in a distance function. Surprisingly, the inputs are varied only rarely within one study, which rules out the possibility of debating the impact of these variables on the efficiency level in detail.

The available budget is the most frequently used input factor category, with 62% of all estimations including it. More precisely, it is employed in 47% of all production functions (Castano and Cabanda, 2007; Sav, 2012l; Sav, 2012k; Fieger, Villano and Cooksey, 2016; Zhang, Bao and Sun,

[20] Aiming for a concise overview, we do not distinguish if the variables are employed in the form of percentages or absolute values.

2016; Agasisti and Belfield, 2017; Bachan, 2017), 100% of all cost functions and 21% of all distance functions (Nemoto and Furumatsu, 2014; Titus, Vamosiu and McClure, 2016).[21] Thus, the budget is usually reflected by the annual overall expenditures of each institution. The studies by Sav (2012k) and Sav (2012l), both employing student service expenditures, are exceptions. Relying on more detailed data, Johnes (2014) distinguishes between academic and administrative costs, Olivares and Wetzel (2014) between operating and personnel costs and Sav (2012a) and Sav (2012b) between academic and student costs. Decomposing the funding of universities further, some authors also include additional budget variables, such as the value of the endowment fund, the costs of outsourcing or other prime costs, taxes and charges (Sav, 2013; Rzadzinski and Sworowska, 2016).

The second most frequently used input factor category concerns the personnel structure, with 83% of all production functions and 98% of all distance functions implementing a related variable. The interpretation of this category is wide-ranging. The overall number of employees, often referring to full-time equivalents, is used in the majority of estimations (Chapple et al., 2005; Castano and Cabanda, 2007; Sav, 2012a; Sav, 2012l; Sav, 2012k; Johnes, 2014; Agasisti, Barra and Zotti, 2016). Extending the overall number, some studies separate teaching and research (Zhang, Bao and Sun, 2016) or academic and administrative staff (Abbott and Doucouliagos, 2009; Sav, 2013; Nemoto and Furumatsu, 2014; Kulshreshtha and Nayak, 2015). In contrast, authors also chose to include only academic (Barra, Lagravinese and Zotti, 2018) or teaching staff (Das and Das, 2014; Laureti, Secondi and Biggeri, 2014; Titus, Vamosiu and McClure, 2016). Slightly more uncommon is the consideration of the number of professors (Miranda, Gramani and Andrade, 2012; Erkoc, 2015; Bolli et al., 2016), the ratio of professors to students (Zoghbi, Rocha and Mattos, 2013), the ratio of staff to students (Bachan, 2017) or the number of tenured and non-tenured positions (Sav, 2012b).

A further input category covers the number of students and it is included in 62% of all production functions and 38% of all distance functions. Again, the interpretation of the input is quite varied. While Das and Das (2014) use the overall number of students and Laureti, Secondi and Biggeri (2014) use the undergraduate enrollment only, most studies separate undergraduate and graduate pupils (Sav, 2012b; Sav, 2012k; Sav, 2012l; Sav, 2013; Zhang, Bao and Sun, 2016; Bachan, 2017). Given that students are seen as an input to the production, some authors also attempt to account for the quality of the attracted students. Johnes (2014), Agasisti, Barra and Zotti (2016) and Barra, Lagravinese and Zotti (2018) use an index in which students are weighted according to their scores in secondary school or their universities entry scores. Similarly, Bachan (2017) directly employs the median entry points and Sav (2012l) and Sav (2012k) use the admission test scores as an input.

[21] We refrain from listing all studies using cost-functions; an overview can be found in Table A.1.

Expanding this approach, some authors also include more detailed student characteristics. The characteristics are thereby evaluated as stand-alone variables in the estimation and are not seen as environmental variables, which is in contrast to other studies (see section 4.5). Thus, the percentage of students that are female, working, domiciled, from state schools, non-white, receiving aid or scholarships and have mothers with HE are used as input factors. Also considered are the age of students and the dropout and retention rates of the institutions (Miranda, Gramani and Andrade, 2012; Sav, 2012a; Sav, 2012b; Sav, 2012k; Sav, 2012l; Zoghbi, Rocha and Mattos, 2013; Bachan, 2017). The study by Zoghbi, Rocha and Mattos (2013) can be seen as an exception to the other studies because it compares three estimations, which differ in the composition of input variables. Using between six and nine different factors, the authors evaluate how more or less information on student characteristics influences the efficiency values. The comparison shows that the efficiency results are robust in regard to the employed input variables. However, given that the authors simultaneously vary the determinants of efficiency, the explanatory power of the results is restricted.

Along with the expenditures as a monetary input, the physical capital plays a role in the production of universities. However, with 47% of all production functions and 4% of all distance functions employing a related variable, it is only occasionally considered. While Castano and Cabanda (2007) and Rzadzinski and Sworowska (2016) focus on the value of property and plants, equipment and consumed materials, the majority of authors chose to focus on the quantity of items. Thus, the number of computers, the computer to student ratio, the numbers of seats in lecture halls and computer laboratories or the library stock (in form of the number of books, etc.) are the employed inputs (Miranda, Gramani and Andrade, 2012; Zoghbi, Rocha and Mattos, 2013; Laureti, Secondi and Biggeri, 2014; Kulshreshtha and Nayak, 2015). Focusing on the role of private philanthropy in the funding of public universities, Sav (2013) considers the value of the art collection.

Given that research is a major part of university activities, 41% of all production functions implement a corresponding input. The included variables are often of a monetary nature, with authors either focusing on research expenditures (Sav, 2012l; Sav, 2012k; Zhang, Bao and Sun, 2016) or research grants (Sav, 2012b; Sav, 2013; Chapple et al., 2005). The inputs used by Zhang, Bao and Sun (2016), who additionally include the value of research equipment, and Chapple et al. (2005), who consider invention disclosures and legal intellectual property spending, are exceptions. The latter study also evaluates how the estimation results change depending on the input used, comparing efficiency values based on either research income or invention disclosures. The authors show that the average technical efficiency is consistent across both evaluations, irrespective of the employed functional form (translog or Cobb-Douglas) and the used output definition (amount or revenue).

Naturally, some studies include further inputs, which do not fit into one of the above categories and are used only once or twice in efficiency evaluations. Showing the diversity of teaching, the number of subjects offered (Das and Das, 2014) and the number of courses taught (Miranda, Gramani and Andrade, 2012; Fieger, Villano and Cooksey, 2016) are such variables that are only irregularly considered. A different approach is taken by Miranda, Gramani and Andrade (2012) and Zoghbi, Rocha and Mattos (2013), who include variables based on students' feedback. The former thereby considers responses on the physical facilities, computer access and library collections of the institutions. The latter, in contrast, uses the feedback to examine the existence of a pedagogical plan.

Output Factors

Similar to the inputs, we grouped the employed outputs into six categories relating to students, graduates, research grants, publications, personal structure and, lastly, other factors. On average, one output factor is used in a production function, four in a cost function and three in a distance function. In comparison to the input, output variables are varied more frequently within one study, which allows us to discuss the influence of these variable on the efficiency level in more detail.

The most frequently used output category relates to the number of students, which is included in 50% of all estimations. More precisely, it is employed in 15% of all production functions, 74% of all cost functions and 55% of all distance functions. Its interpretation is, similar to its usage as an input, quite varied. Only a few authors include the absolute number of students (Miranda, Gramani and Andrade, 2012; Kulshreshtha and Nayak, 2015; Bolli et al., 2016; Rzadzinski and Sworowska, 2016). Most authors divide the overall amount, relying on arguments concerning the duration or cost of study. The separation according to the level of education is widespread, splitting undergraduate, graduate and postgraduate students (Robst, 2001; Mensah and Werner, 2003; Fu, Huang and Tien, 2008; Johnes, Camanho and Portela, 2008; Kuo and Ho, 2008; Abbott and Doucouliagos, 2009; Mamun, 2011; Sav, 2012c; Sav, 2012d; Sav, 2012f; Sav, 2012h; Sav, 2012i; Sav, 2012j; Agasisti and Haelermans, 2016; Titus, Vamosiu and McClure, 2016). Likewise, the division of students based on the subject groups is popular, distinguishing mainly between science and non-science and, in some instances, medicine or additional fields (Agasisti and Johnes, 2010; Johnes and Schwarzenberger, 2011; Olivares and Wetzel, 2014; Agasisti, 2016; Gralka, 2018). Irrespective of the chosen separation practice concerning students, PhD students are frequently included as an additional output (Johnes, 1998; McMillan and Chan, 2006; Johnes, Camanho and Portela, 2008; Johnes and Salas-Velasco, 2007; Agasisti and Johnes, 2010; Johnes and Schwarzenberger, 2011; Agasisti and Haelermans, 2016). [22] Relying on more extensive datasets,

[22] The distinction between undergraduate, graduate and postgraduate students differs across studies. If PhD students are listed as a separate output, we evaluate them as such.

some studies include both forms of division and split students according to the level of education as well as the subject group (Izadi et al., 2002; Stevens, 2005; McMillan and Chan, 2006; Horne and Hu, 2008; Johnes, Johnes and Thanassoulis, 2008; Lenton, 2008; Johnes and Johnes, 2009; Nemoto and Furumatsu, 2014; Johnes and Johnes, 2016). Likewise, exploiting a comprehensive dataset, but choosing a different approach, Lu and Chen (2013) modify the output through additional quality variables. They adjust the number of students, of students attending extended education and of acquired certificates, by corresponding ratios. These ratios include the share of students to teacher, professors to all teachers, extended students to all students and higher level certificates to all certificates. Also aiming to account for the quality of teaching, two studies in our sample take the study progress of students into account. Agasisti (2016) thereby focuses on students that are enrolled below the regular duration of the course. Contrasting the results to an estimation based on the total number of students, he shows that the results change slightly, but without "dramatic differences" (Agasisti, 2016, p. 61). However, looking at the marginal costs, he concludes that if more students were able to stay within the legally set timeframe, unit costs would drop significantly. Taking a further step, Laureti, Secondi and Biggeri (2014) calculate an output variable that represents full-credit-equivalent students. The measure is based on the ratio between the total amount of credits achieved by all students and the theoretical maximum number of credits that, according to the agenda, should be obtained. They compare the resulting efficiency values to the estimation values based on regular graduates (see section below). Similar to the corresponding input category, authors moreover aim to take the quality of education or the characteristics of students into account as an output. While only two studies control for the quality of the incoming students (Stevens, 2005; Lenton, 2008), more authors aim to account for the quality of education at the institution. The frequently used variables include the average exam grades, the graduation rate or the percentage of good degrees awarded (Mensah and Werner, 2003; Sav, 2012b; Sav, 2012k; Zoghbi, Rocha and Mattos, 2013; Sav, 2016; Bachan, 2017). Evaluating the characteristics of students, some authors focus on the minorities by including variables such as the percentage of students with low income grants or non-white student enrollment (Sav, 2011; Sav, 2012c; Sav, 2012d; Sav, 2012e).

The second most frequently used output category refers to graduates. It is employed in 30% of all production functions, 16% of all cost functions and 38% of all distance functions. Again, some authors employ the overall number (Kempkes and Pohl, 2008; Kempkes and Pohl, 2010; Rzadzinski and Sworowska, 2016), while others prefer to differentiate graduates according to degrees or subject groups. Thereby, the separation of undergraduate, graduate and postgraduate students (Worthington and Higgs, 2011; Agasisti and Johnes, 2015; Agasisti and Haelermans, 2016) is more frequent than the separation of subjects, which is only done by Agasisti (2016).

Besides, the weighting of graduates according to their degree is common (Das and Das, 2014; Johnes, 2014; Agasisti, Barra and Zotti, 2016; Barra, Lagravinese and Zotti, 2018). Also recurrently employed is the number of regular graduates, referring to those that finish within the regular duration of the course (Laureti, Secondi, and Biggeri, 2014; Agasisti, 2016; Agasisti and Belfield, 2017). Once more, studies that compare different graduation output measures are informative, as is done by four studies in our sample. Similar to his evaluation based on students, Agasisti (2016) additionally tests if the overall number of graduates leads to a similar efficiency assessment as the number of graduates that are within the regular duration of the course. He shows that, again, the results differ only slightly between the two indicators. Aiming for a comprehensive sensitivity check Agasisti and Belfield (2017) contrast the results of their baseline efficiency estimation, based on the absolute number of awarded certificates, to the results of other performance indicators. As alternative outputs, they use the weighted count of given credits (favoring credits of graduates), the number of associated degrees granted and the number of graduates that finish within 150% of the regular course time. While the results of all four estimations are positively correlated and show a similar trend, the authors emphasize that they are not identical and, in particular, the findings on the environmental variables change with the employed output. Likewise, varying the output variable but focusing on the gender of students, Sav (2012l) separately estimates efficiency based on all, on male or on female graduates. He confirms the results from Agasisti and Belfield (2017), showing that the parameters and the efficiency values differ slightly with the chosen output. In contrast, focusing on the subject group, Sriboonchitta (2012) compares efficiency evaluations based on non-science or science graduates. However, since both outputs are contrasted to the total cost of producing graduates, the informative value of the comparison is limited.

Concluding the evaluation of both outputs categories that represent the teaching activities of institutions, it is useful to examine studies that contrast the two variables to see if they lead to a similar assessment. Aiming to avoid double counting, no study in our sample implements both outputs within an estimation, but four studies contrast the results of the two indicators. Agasisti and Haelermans (2016) provide the most prominent study on that topic, addressing the choice in their research question. They contrast evaluations based on the number of students ("cost for activity model") and on the number of graduates ("cost for performance model") for two countries. Depending on the chosen teaching output, one or the other country is relatively more efficient. They, therefore, summarize that the selection has a strong influence on the resulting efficiency values[23]. Using a similar approach, but evaluating only one country, Agasisti (2016) finds that the

[23] In their study, Agasisti and Gralka (2017) likewise contrast the two indicators in a cross-country evaluation and show that they lead to overall similar results. Since the analysis is only available as a working paper, it is not included in the present survey.

distribution of efficiency scores is very different when students instead of graduates are considered. More universities are classified as efficient if the number of students is employed than if graduates are used. This result is confirmed by Laureti, Secondi, and Biggeri (2014), who contrast the results based on their measure of full-credit-equivalent students with an evaluation based on regular graduates. They show that the inclusion of students leads to an overall higher efficiency assessment. However, this result could be limited to the Italian HE sector, since the latter two studies are performed for this country. Only Rzadzinski and Sworowska (2016) evaluate the subject for a different country and confirm that the inclusion of students leads to higher efficiency values for Polish institutions. Unfortunately, the latter study merely compares the range of efficiency values. It forgoes the opportunity to show the distribution or the rank correlation of the efficiency values in depth. Hence, we recommend that future researchers discuss their output choice in more detail and include such comparisons.

Moving from the teaching to the research activity of universities, the sample of studies displays that research grants are the overall most frequently used indicator for research output. The variable is employed in 86% of all cost functions and 48% of all distance functions. It is not used as an output in any production function, presumably since only one output can be evaluated in that type of function. As briefly mentioned above, the usage of research grants as an output is strongly debated in the literature. Authors employing the variable argue that it reflects the market value of conducted research and can, therefore, even be considered a quality adjusted proxy for output. An advantage is also the easy access to accurate data, which is often collected by the statistical offices. Nevertheless, critics argue that it is not a clear cut output measure, since the funds are not only spent on research but also on other facilities, which are an input for production. In addition, the funds are commonly distributed unequally over a subject group. Within the present paper, we refrain from discussing the pros and cons of this (and the below considered) research indicator in more detail, accepting that it is a long-standing discussion in the literature.[24] However, in light of the prominent critique concerning the usage of research grants and given its regular usage, it is surprising that the literature refrained from introducing adjustments to this indicator. The majority of studies use the overall value of research grants (Johnes, 1998; Robst, 2001; Izadi et al., 2002; Mensah and Werner, 2003; Stevens, 2005; McMillan and Chan, 2006; Johnes and Salas-Velasco, 2007; Johnes, Johnes and Thanassoulis, 2008; Kempkes and Pohl, 2008; Kuo and Ho, 2008; Lenton, 2008; Johnes and Johnes, 2009; Agasisti and Johnes, 2010; Kempkes and Pohl, 2010; Johnes and Schwarzenberger, 2011; Mamun, 2011; Sav, 2011; Worthington and Higgs, 2011; Sav, 2012c; Sav, 2012d; Sav, 2012e; Sav, 2012f; Sav, 2012g; Sav, 2012i; Sav, 2012j; Johnes, 2014; Nemoto and Furumatsu, 2014; Agasisti and Johnes, 2015; Agasisti, 2016; Agasisti, Barra and Zotti,

[24] See Gralka, Wohlrabe and Bornmann (2018) for a discussion on this debate.

2016; Agasisti and Haelermans, 2016; Johnes and Johnes, 2016; Sav, 2016; Titus, Vamosiu and McClure, 2016; Barra, Lagravinese and Zotti, 2018). Only the studies by Olivares and Wetzel (2014) and Gralka (2018) separate the overall amount according to science and non-science subjects.

As an alternative to research grants, publications are frequently used to measure the research output of institutions. The indicator has the advantage that subject groups can be more accurately reflected and a quality weighting, for example, through citations, is possible. Nevertheless, it is a retrospective measure, a variety of publication outlets exist and accurate data is difficult to obtain. The indicator is employed in 17% of all production functions, 5% of all cost functions and 52% of all distance functions. The high share for the latter type of function is in part related to the publication period. Due to the difficulty of obtaining representative data on the institutional level, publications are primarily employed in recent studies. The study by Johnes, Camanho and Portela (2008) is the first to use this indicator in our sample.[25] Since distance functions are likewise part of a more recent development (see section 4.2), a relatively high percentage is to be expected. The majority of studies employing publications as a research indicator refer to the absolute number (Johnes, Camanho and Portela, 2008; Abbott and Doucouliagos, 2009; Lu and Chen, 2013; Bolli et al., 2016). It is worth noting that most studies do not provide information on what type of publications are included in the measure or how they are aggregated from the author to the institutional level. The exceptions are often studies that separate different types of publications in their efficiency evaluation, thereby providing more information on the output. Within the studies in our sample, authors distinguish between books, book chapters, journals and other publications forms (Abbott and Doucouliagos, 2009; Worthington and Higgs, 2011; Kulshreshtha and Nayak, 2015). Aiming to account for the quality of publications, Erkoc (2015) separately includes the overall number of publications, the number of publications in SSCI indexed journals and the number of citations during the last four years. Mentionable is the study by Zhang, Bao and Sun (2016), since it contrasts five efficiency evaluations. The authors use either the total number of publications, the number of science or non-science publications or the number of publications in domestic or international journals as outputs. However, since the employed input factors are not split according to the same grouping, the resulting efficiency values can be interpreted only to a limited extend.

Concluding the examination of both research outputs, it is again of interest to examine if any studies contrast the two variables, evaluating if they lead to a similar assessment. However, no study in

[25] In fact, Johnes and Johnes (1993) were the first to employ publications as a research output in an efficiency evaluation, using DEA. Later, Warning (2007) implemented the indicator within a SFA. The latter analysis is not included in the present sample since it is published as part of a book.

our sample separately evaluates efficiency for both research indicators.[26] Only the study by Worthington and Higgs (2011) employs research grants and publications in their estimation of efficiency, but it refrains from contrasting the two and instead includes them both as outputs in the estimation.

A fifth output measure concerns the personal structure of institutions. The output is employed in 17% of all cost functions and 7% of all distance functions. Interestingly, the majority of studies avoid the inclusion of an absolute number. The usage of ratios (foremost the student to teacher ratio) or percentage data, as the percentage of faculty with tenure, is more common, (Fu, Huang and Tien, 2008, Lenton, 2008; Sav, 2011; Sav, 2012c; Sav, 2016; Titus, Vamosiu and McClure, 2016). Exceptions are the studies by Sav (2012e), who additionally includes absolute numbers (separating the faculty according to the length of their contract) and Mamun (2011), who includes the number of teaching and non-teaching staff. Evaluating the impact that the choice of variables can have on the efficiency results is again of interest. Within out sample, Lenton (2008) is the only study that compares estimations that differ in the usage of variables concerning the personal structure.[27] The author includes percentage data on personal staff in the first estimation specification and leaves it out of the second. However, the regression results from these first two specifications are (deliberately) only discussed in light of their significance. Based on the comparison, the variables concerning the personal structures are deemed valid and are then included in the third specification. The efficiency values are discussed only for the latter specification.

In addition to the above listed outputs, some studies include further outputs that do not fit into one of the above categories. A handful of studies employ the number of teaching hours per institution in their efficiency evaluation (Fieger, Villano and Cooksey, 2016; Sav, 2011; Sav, 2012a; Sav, 2012e; Sav, 2012g; Sav, 2016). A second group of further outputs depicts revenues, with some authors even differentiating between different types of revenue sources (McMillan and Chan, 2006; Castano and Cabanda, 2007; Johnes, Johnes and Thanassoulis, 2008; Mamun, 2011; Sav, 2011; Sav, 2012d; Sav, 2013; Rzadzinski and Sworowska, 2016). Further unique outputs are the number of licenses registered (Chapple et al., 2005), the age of the institution (Mamun, 2011), the assets of the university (Mensah and Werner, 2003; Sav, 2016) and the ratio of students per square meter of building space (Fu, Huang and Tien, 2008).

[26] This research gap was filled by Gralka, Wohlrabe and Bornmann (2018) by contrasting efficiency values based on research grants or publications. The study is not part of the present sample since it was published after 2017.

[27] The studies by Sav (2012c) and Sav (2012d) should also be mentioned in this context. Both studies include information on the personal structure, either as an output or a determinant of efficiency (for a discussion, see section 4.5)

Looking at the considered output factors it has to be noted that the amount of studies that included the third mission of the institution is surprisingly low. Within our sample, only the study by Chapple et al. (2005) can be sorted into this area of research, given that it includes recorded licenses in the efficiency analysis.[28] We, therefore, hope that future researchers will turn toward this subject.

Prices

Along with the enquiry, if the inputs and outputs are fully utilized (technical efficiency), researchers are interested in evaluating whether the observed combination of factors is the best (allocative efficiency). Prices, therefore, play an important role as an assessment criterion. It is in general assumed that the prices are exogenous and efficiency is obtained by a choice on the level of input.[29]

Input prices are used in 70 estimations within 26 studies in our sample,[30] among which, 46 estimations include one price, 18 estimations include two prices and six estimations even include three prices. The most frequently used variable is wage, which is employed in all estimations that include at least one price variable. Wage is thereby commonly represented by the average annual wage rate over all employees, which is approximated by the ratio of total personnel expenditures to the number of employees (Stevens, 2005; McMillan and Chan, 2006, Fu, Huang and Tien, 2008; Kuo and Ho, 2008; Kempkes and Pohl, 2008; Kempkes and Pohl, 2010; Mamun, 2011; Sav, 2011; Sav, 2012a; Sav, 2012b; Sav, 2012c; Sav, 2012d; Sav, 2012h; Sav, 2012j; Sav, 2012l; Lu and Chen, 2013; Johnes and Johnes, 2016; Titus, Vamosiu and McClure, 2016; Gralka, 2018). However, some studies also employ more accurate information on the average teacher pay (Lenton, 2008) or the compensation paid to teachers (Robst, 2001; Lenton, 2008). Five studies in our sample include two wage variables, separating the price of academic and non-academic labor (Worthington and Higgs, 2011) or separating the average salary according to the length of contracts (Sav, 2012e; Sav, 2012f; Sav, 2012g; Sav, 2012i). In addition to the wage, eight studies include prices to reflect other inputs, of which, the price of capital is foremost. The latter variable is measured in a variety of ways, including the ratio of capital cost to the area of the schoolyard or the overall value of all buildings at the end of the year (Fu, Huang and Tien, 2008; Worthington and Higgs, 2011; Sav, 2011; Sav, 2012a; Sav, 2012f; Sav, 2012h; Sav, 2012i; Lu and Chen, 2013). In addition, the teaching price (teaching cost divided by the number of students, Lu and Chen, 2013) and the value of equipment

[28] Hu, Yang and Chen, 2014 use the number of registered patents in their SFA estimation. The study is not part of the present sample since the authors estimate efficiency on the country (instead of institutional) level.

[29] The applied function has to be homogenous in the input prices (Kumbhakar, Wang and Horncastle, 2015). Various approaches to ensure the price homogeneity condition exist. A straightforward and frequently used solution is the arbitrary choice of one price as the normalizing price. We refrain from discussing the approaches taken in the evaluated studies further.

[30] Since output prices are not employed in the present sample, they are not discussed further.

and libraries (Sav, 2012a; Sav, 2012i) are the employed price variables. No study in our sample contrasts an efficiency estimation that varies in the inclusion of prices.

As stated in section 4.2, prices can be included if a cost function or distance function (Coelli et al., 2005) is used for the estimation. It is therefore surprising that in three studies in our sample a production function, which includes at least one input price, is used for the evaluation (Sav, 2012a; Sav, 2012b; Sav, 2012l). Notably, the majority of studies looking at HE institutions do not include prices for the efficiency evaluation, even when a cost or distance function is employed. This could be driven partly by the available data and partly by the assumption that input prices in the HE sector do not vary much, which is consistent with the assumption of a competitive market.

Dummies

Dummy variables are used in 67 estimations within 19 studies in our sample. The number of implemented factors thereby varies between one up to seven, with one estimation as an outlier even incorporating 50 dummy variables (reflecting US states, Agasisti and Belfield, 2017). The binary variable is most commonly used to reflect the offered subjects or the location of institutions. Arguing that medicine and engineering subjects bring about higher costs, 21 estimations include a dummy for the existence of a medicine faculty (Kempkes and Pohl, 2008; Agasisti and Johnes, 2010; Kempkes and Pohl, 2010; Sav, 2012c; Sav, 2012d; Lu and Chen, 2013; Agasisti and Johnes, 2015; Titus, Vamosiu and McClure, 2016) and eight estimations for the presence of an engineering subject (Kuo and Ho, 2008; Kempkes and Pohl, 2008; Agasisti and Johnes, 2010; Lu and Chen, 2013). Similarly, Mensah and Werner (2003) emphasize art institutions and Kuo and Ho (2008) include a factor that represents the diversity of academic fields (measured by the number of academic departments in one field relative to the overall number). Focusing on the location of the institutions, studies likewise use dummy variables to differentiate between countries (Agasisti and Haelermans, 2016; Bolli et al., 2016), states (Agasisti and Belfield, 2017) or regions (Lenton, 2008; Zoghbi, Rocha and Mattos, 2013; Agasisti, 2016; Gralka, 2018). Also incorporating the location of institutions, Titus, Vamosiu and McClure (2016) examine if they are located in an urban area. Furthermore, dummies are used to reflect the offered degrees, marking if a university has a graduate or PhD program (Robst, 2001; McMillan and Chan, 2006; Kuo and Ho, 2008; Lenton, 2008; Sav, 2012h; Titus, Vamosiu, and McClure, 2016). Additionally, they are employed to distinguish between different types of institutions, marking private (Lu and Chen, 2013; Agasisti and Johnes, 2015), research (Robst, 2001; Kuo and Ho, 2008) regional (Mensah and Werner, 2003) or historically black institutions (Titus, Vamosiu and McClure, 2016). Johnes (2014) and Bachan (2017) use a UK specific dummy variable, considering the statuses of the institutions before the UK Higher Education Act of 1992.

Two studies in our sample contrast evaluations, which include or exclude the employed dummies. The most detailed study in that regard is published by Kempkes and Pohl (2010). The authors use dummies to control for the faculty composition of institutions, employing variables for engineering and medical faculties. By contrasting the estimation results to a model without dummies, they show that the dummies are significant and that their usage is supported by a performed likelihood ratio test. In addition, the authors point out that some of the interaction terms of the output variables and dummies are significant, "[…] indicating that institutions with medical and/or engineering faculties not only have different cost levels but also different marginal cost structures." (Kempkes and Pohl, 2010, p. 2071). The authors emphasize the necessity to control for the faculty composition, stating that a non-consideration will lead to biased results. Based on that reasoning, they refrain from showing the efficiency values for the model without dummies. In a similar manner, but evaluating more than one country, Bolli et al. (2016) estimate efficiency based on a model with and without country dummies. The authors show that the regression results are robust in regard to the usage of these dummies. However, the efficiency values are again not compared.

While dummy variables are repeatedly included in efficiency analysis, the reasoning behind and their possible implications are rarely discussed. Especially in an evaluation of the HE sector, one should discuss why the identified groups supposedly produce in a different way or exhibit different technologies. Unfortunately, we are not aware of a study discussing the general implementation of dummy variables within efficiency analysis in a structured and detailed manner.

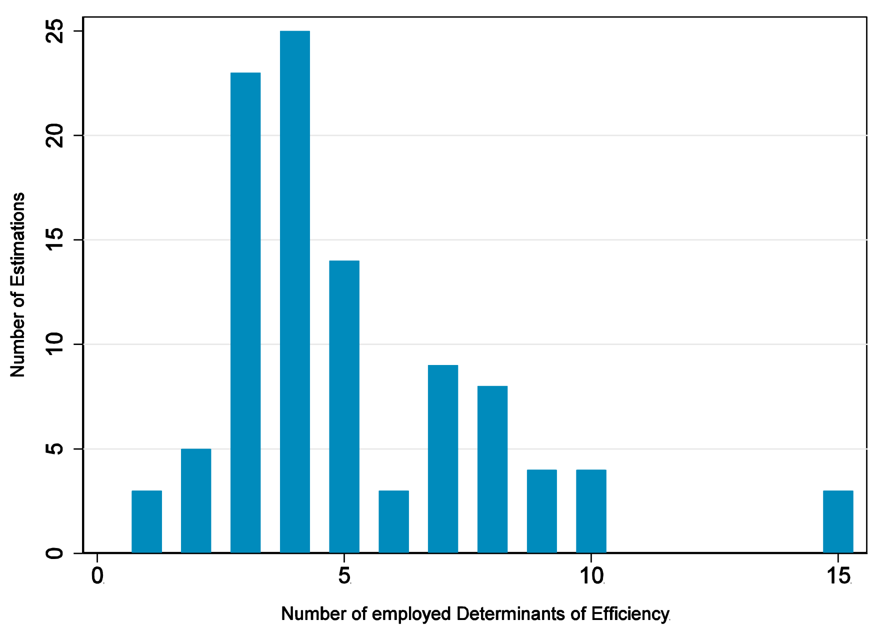
Examining the categories of utilized input and output variables, it is obvious that there is a core composition of variables representing the expenditures as well as the teaching and research tasks of universities. However, the choice of variables complementing those is diverse. Surprisingly, authors largely refrain from varying the employed factors, especially the considered inputs, within one study. Unfortunately, this limits the possibility to debate the impact of the chosen factors on the efficiency level further. From studies that do vary the variables, it is observable that the efficiency values differ slightly with the composition. However, as Agasisti and Belfield (2017) state, "This result is not worrying per se; it is reasonable that efficiency is not an absolute concept, and is instead dependent upon the specific output included in the empirical analysis." (Agasisti and Belfield, 2017, p. 251). The evaluation of factors moreover showed that the amount of studies that include the third mission of institution is surprisingly low. Additionally, most studies refrain from including prices and dummy variables in the estimation. We, therefore, recommend that future researchers discuss their choice of variables more strongly, compare different compositions of factors and look at the third mission activities. When aiming to compile an overview of estimations, we strongly emphasize the need to take the type of function into account.

4.5 Determinants of efficiency in education

In the present section, focus is laid on the determinants of universities efficiency. Following the argument by Coelli et al. (2005), the chosen variables should thereby neither be inputs nor outputs of the production process, should be outside of the institutions control and (nonetheless) should influence the producer's performance. The subsequent description focuses on the estimation level again, since determinants are often varied within one study. Information on the related SFA estimation specifications can be found in section 4.3. We deliberately keep the following paragraph short, given that De Witte and López-Torres (2017) provide a recent and well-structured overview of employed environmental variables. While the authors mainly focus on studies employing DEA, we believe that similar determinants of efficiency can be employed within the DEA and SFA, given that both methods possess similar methodical restrictions on this aspect (contrary to the chosen input and output factors). We show that the literature employs a wide variety of determinants. However, a core model, which is based on a theoretical foundation, seems to be missing.

Within our sample, determinants of efficiency are employed in 101 estimations within 25 studies. The quantity of employed determinants thereby varies between one up to fifteen, with the majority of studies implementing four in one estimation (see Figure 4). We grouped the employed determinants into eight categories relating to students, personal structure, budget, institutions, subject specialization, region, quality of education and research. The categories are presented according to the frequency of their usage. Notably, a broad variety of determinants of efficiency are used. Even the variables, which represent similar influencing factors, vary in minor

Figure 4: Number of Employed Determinants of Efficiency (Count of Estimations)



(calculation) details. Aiming for a concise overview, we, therefore, refrain from displaying the share of each determinant of the total, as we did for the inputs and outputs.

Predictably, the determinants of efficiency are most frequently used to evaluate the students that are registered in the evaluated institutions. Authors take account of the total enrollment as well as the change in enrollment from the year before (Robst, 2001; McMillan and Chan, 2006; Sav, 2012d; Sav, 2012f; Sav, 2012i; Sav, 2012j; Laureti, Secondi and Biggeri, 2014; Fieger, Villano and Cooksey, 2016). As an exception to other studies, Laureti, Secondi and Biggeri (2014) thereby focus on the freshmen enrollment and additionally take the pupils marks from and the subject specialization of the secondary school into account. Relating to the enrollment, authors also control for the origin of students, distinguishing between native and overseas, European and non-European or regional and national students (Stevens, 2005; Abbott and Doucouliagos, 2009; Laureti, Secondi and Biggeri, 2014). In addition to the origin, the inclusion of student's characteristics as determinants of efficiency is common. Thus, some researchers concentrate on the ethnic background, including variables such as the share of non-white, African American or Hispanic students (Stevens, 2005; Sav, 2012e; Zoghbi, Rocha and Mattos, 2013; Agasisti and Belfield, 2017), while others, in a broader sense, include the share of minority enrollment (Sav, 2012c; Sav, 2012f; Sav, 2012i). Likewise, controlling for the background of the students, the share of students that are from lower classes, receiving low income grants or have English as a second language are included as factors (Stevens, 2005; Sav, 2011; Sav, 2012e; Sav, 2012f; Sav, 2012g; Sav, 2012h; Fieger, Villano and Cooksey, 2016; Sav, 2016). Taking a further step, Zoghbi, Rocha and Mattos (2013) evaluate how the maternal education of students influences the efficiency of institution. Focusing on the composition of students, authors also include the age of students (Zoghbi, Rocha and Mattos, 2013; Laureti, Secondi and Biggeri, 2014; Fieger, Villano and Cooksey, 2016; Agasisti and Belfield, 2017) as well as the gender (Stevens, 2005; Zoghbi, Rocha and Mattos, 2013; Laureti, Secondi and Biggeri, 2014; Barra, Lagravinese and Zotti, 2018). Other factors relating to the student structure are the share of part-time students (McMillan and Chan, 2006; Agasisti and Belfield, 2017), working students (Zoghbi, Rocha and Mattos, 2013) or disabled students (Fieger, Villano and Cooksey, 2016).

The second most frequently used category of determinants concerns the personnel structure. Authors include variables that reflect details on the issued contracts, as the amount of faculty employed in tenure and non-tenure track positions, the length of contracts and the average wage (Sav, 2011; Sav, 2012c; Sav, 2012d; Sav, 2012e; Sav, 2012g; Sav, 2012i; Sav, 2012j; Sav, 2012k; Sav, 2012l). Information on the job level, with determinants representing the amount of academic staff, executive staff, apprentices and trainees, professors and senior staff are also incorporated (Stevens, 2005; Abbott and Doucouliagos, 2009; Sav, 2013; Fieger, Villano and Cooksey, 2016).

The determinants of efficiency are likewise used to reflect characteristics of the employees, focusing on the female staff, non-white staff or the age of employees (Stevens, 2005; Sav, 2012l). The ratio of students to faculty is also contained within this category and is used by two studies in our sample (Sav, 2012c; Sav, 2012d).

A third category depicts details on the budget of institutions. The determinants of efficiency are thereby mainly used to account for the composition of funds and revenues. Authors consider the universities international budget shares (Bolli et al., 2016), liabilities (Sav, 2012f; Sav, 2012h), private (Sav, 2012h; Bolli et al., 2016; Sav, 2016; Barra, Lagravinese and Zotti, 2018) or state funds (Robst, 2001; Mensah and Werner, 2003; Kuo and Ho, 2008; Mamun, 2011; Sav, 2011; Sav, 2012i; Sav, 2012j; Sav, 2016), tuition fees (Agasisti, Barra and Zotti, 2016; Bolli et al., 2016; Barra, Lagravinese and Zotti, 2018) and administrative or operating expenses (Mensah and Werner, 2003). Besides the composition, some studies also aim to include the change of available assets. Hence, the lags of budget shares (Bolli et al., 2016) as well as the change in total revenues (McMillan and Chan, 2006) are used as determinants of efficiency.

In addition to the budget, some further institutional variables are included as determinants of efficiency. Most researchers thereby aim to evaluate how the type of institution influences its productivity. Hence, applied universities, research institutions, vocational colleges and Dawkins institutions (Australia specific) are identified (Robst, 2001; Kempkes and Pohl, 2008; Abbott and Doucouliagos, 2009; Agasisti and Belfield, 2017). Likewise, popular is the evaluation of institutions that award doctoral or master degrees (Robst, 2001; Sav, 2012k; Sav, 2013). Taking a further step, Mensah and Werner (2003) make use of the Carnegie foundation classification, which describes institution diversity in the US, for their determinants of efficiency. The factors employed also include the age of the university (Chapple et al., 2005; Castano and Cabanda, 2007; Mamun, 2011; Barra, Lagravinese and Zotti, 2018) and its ownership (Castano and Cabanda, 2007; Sav, 2012f; Zoghbi, Rocha and Mattos, 2013). In addition, Castano and Cabanda (2007) assess how the degree of autonomy influences efficiency, while Kempkes and Pohl (2008) consider if the university operates under comparatively liberal or restrictive state regulations. Somewhat exceptional in that category is one of the determinants used by McMillan and Chan (2006), who include the proportion of classes with less than 26 students.

Within the institutional context, there are also determinants that depict the subject specializations of institutions. Thereby, the assessment regarding if a medical school and hospital have an influence on the efficiency of the universities is most common (Chapple et al., 2005; Sav, 2013; Barra, Lagravinese and Zotti, 2018). However, researchers also look at other fields of study, evaluating, for example, the overall number of programs offered or the specialization among programs (Stevens, 2005; McMillan and Chan, 2006; Abbott and Doucouliagos, 2009; Laureti, Secondi and Biggeri, 2014).

A further category of determinants concerns the region the institutions are situated in. While only one study in our sample takes the actual location into account (Kempkes and Pohl, 2008), another analysis uses a different definition and assess if the remoteness plays a role in the efficiency evaluation (Fieger, Villano and Cooksey, 2016). Most authors, however, choose to include the gross domestic product of the state or region (Chapple et al., 2005; Kempkes and Pohl, 2010; Zoghbi, Rocha and Mattos, 2013; Laureti, Secondi and Biggeri, 2014; Agasisti, Barra and Zotti, 2016; Barra, Lagravinese and Zotti, 2018). Likewise, considering variables at the state level, Zoghbi, Rocha and Mattos (2013) include the average years of schooling, and Kempkes and Pohl (2008) include the share of the population aged 18–35. To capture competition between universities, Agasisti, Barra and Zotti (2016) include the market share of the universities, measured as the ratio between the enrollment number at the university and the total number of enrollment in the region.

The quality of education is difficult to quantify and, hence, is only rarely included as an output in efficiency evaluations (see section 4.4). Studies that include the related determinants of efficiency are, therefore, of particular interest. Authors thereby occasionally include graduation rates (McMillan and Chan, 2006; Sav, 2012g; Fieger, Villano and Cooksey, 2016) and the dropout rates (Zoghbi, Rocha and Mattos, 2013). However, one study also takes the average final degree mark (Laureti, Secondi and Biggeri, 2014) and another takes the share of first-class and upper-second-class degrees, referring to a United Kingdom specific degree classification, into account (Stevens, 2005).

Alongside teaching, research can be seen as the main activity of an institution. It is, therefore, surprising, that only two studies include a related determinant of efficiency. Chapple et al. (2005) include a variable that depicts the research intensity, while Robst (2001) includes the amount of research expenditures.

Concluding the examination of the employed determinants, two aspects have to be noted critically. First, the choice of determinants often seems to be data driven, with a clear argumentation for the implementation of the variables frequently missing. Even though we refrain from reporting the reasoning behind the determinants in the present paper, we think that some of the above noted determinants display this shortcoming even without further exposition. Second, opposed to the above given definition by Coelli et al. (2005), the examination reveals that some of the chosen determinants are either known as input or output factors or are arguably within the institutions control. While this seems to be known to some of the authors, who make related statements in their argumentation, the determinants are used, nonetheless. In this context, the examinations by Sav

(2012c) and Sav (2012d) are worth mentioning. In both studies, efficiency estimations are contrasted, in which information is included either as an output or as a determinant of efficiency. In a similar manner, McMillan and Chan (2006) vary between including their chosen determinants in a second stage regression or as environmental variables (see section 4.3 for methodical details). We hope that these observations stimulate critical thinking and provoke a debate on the choice and suitability of determinants.

The review of determinants shows that the literature employs a wide variety of determinants. In contrast to the employed input and output factors, there seems to be no core model for determinants of efficiency so far. While the teaching side of universities is analyzed quite frequently through students or quality related variables, only two studies include further determinants concerning the research aspect of institutions. Further variables mainly concern institutional characteristics, such as the personal structure, the budget, the subject specialization and the region the institution is located in. It has to be critically noted that the choice of determinants often seems data driven, and a clear argument for the implementation of the variables is frequently missing. Moreover, some of the chosen determinants are (arguably) within the control of institutions and should, therefore, be classified as input or outputs. Given the wide variety and obvious controversy, we think that a debate on the choice and suitability of determinants and a resulting clear recommendation would be truly beneficial for future research.

5. Discussion

Assessing the relative performance of institutions is a daunting task, given the diversity of situations and environments. For the calculation, researchers have to make simplifying assumptions, which concern the underlying method as well as the chosen variables. Hence, it is crucial that authors specify and argue their respective choice. Based on the above review, the following paragraph highlights some remaining potential. We begin by summarizing some shortcomings of the reviewed literature on HE institutions. To complete the discussion, we also address the limitations of the present paper.

The above review shows that the literature is not without issues. A major concern that became obvious in the examination of the underlying assumptions (section 4.2) and applied specifications (section 4.3) relates to omitted information. Occasionally, authors forget to state which underlying functional form or which distribution of the efficiency term they assumed. Even worse, some of the reviewed studies did not specify which specification of the SFA was used for the estimation. The issue continues in the description of the utilized factors (section 4.4), with authors occasionally failing to provide detailed information on the variables. The most prominent example is the output factor publications, where almost no study in our sample provides information on the types of publications counted or how they were aggregate to the institutional level. The problem of missing information also relates to the often absent discussion regarding the respective choice of the determinates of efficiency (section 4.5). Since missing information and argumentation are avoidable shortcomings, we hope that future authors ensure that all necessary information is provided. The structure of the review can serve as a rough guideline.

Another aspect that is in need of improvement concerns the presentation of the estimation results. Examining the literature, it is obvious that there is no standard that specifies which information should be given and how it should be presented. While most authors present the regression table either in the text or appendix (within our sample 96%), the presentation of the efficiency values varies greatly. Efficiency values are either provided in the form of the mean for the whole sample (often as descriptive statistics), for each institution (often as a table in the appendix) or as a graph (often as a histogram). While most studies choose the first version and provide a descriptive statistic, the perimeters differ. Statistics are either shown for the overall sample or subgroups, for the whole timeframe or selected years, including or excluding minimum or maximum values. Similar options of display are possible for second and third versions, with authors showing values either for the whole timeframe or for selected years. While it is not the aim of the present paper to specify which version of display is the best, we want to emphasize that there is a discrepancy and a resulting need for a standard. Only if studies on HE institutions start to provide the efficiency values in a standardized way, is it possible to compare them. It would then be feasible to perform a meta-regression analysis, as was done by Bravo-Ureta et al. (2007) for the farming sector or by Aiello and Bonanno (2016) for the banking area. The analysis would allow to retrace the causes for heterogeneity in the efficiency results between studies. The reviewer could differentiate between the variations in efficiency values that are due to the underlying sample, the assumptions, the applied specifications and the utilized factors. If future authors are in doubt about how much information to present, we would suggest providing all efficiency values (for each institution and year) in the appendix. Should researchers deliberately decide against displaying institutional efficiency values, they should not forget to mention the considered institution by name.

In addition, the overview of applied specifications (section 4.3) and utilized factors (section 4.4) showed that further potential exists for conducting a sensitivity analysis. While some authors, such as Eagan and Titus (2016), already emphasize this aspect, more comparisons of different estimations would be beneficial. In particular, the chosen input factors are rarely contrasted within one study. The additional examination would improve the studies themselves, showing if the provided results are robust to changes in assumptions, specification and factors but are also of interest to the research community, allowing further statements regarding the sensitivity and dependence of efficiency values.

The present paper also has its limitations. Creating a review of papers is a difficult undertaking. Studies are written in various contexts and use different approaches. In addition, the bundling of inputs, outputs and determinants into categories is challenging. We have, therefore, chosen a straight forward approach for the present review. The contents of the studies are presented in a type of "accounting approach", aiming to show the reader what has been done and where to look. It would have been equally interesting to take a closer look at the reasoning of the authors, evaluating if and how they explain their assumptions and choices. Likewise, the closer assessment of connections would be revealing, observing, for example, if certain factors are frequently used in distinct combinations or only with certain specifications. As it is, this remains open for future research. In addition, a review is only as good as the underlying sample of studies. Aiming to provide a concise overview, we had to introduce some restrictions to our sample procedure (see section 3.1). Hence, we only considered studies that were published in peer-reviewed journals before or in the year 2017. Realizing that this could lead to a slightly biased sample, we additionally included footnotes at the appropriate sections that refer to excluded, but relevant, studies and current developments. Despite our comprehensive search strategy (see section 3.2), we cannot rule out that some relevant studies may have been missed.[31]

6. Conclusion and Outlook

For the present paper, we have taken a step back to assess where we currently stand and what we know. To this end, we systematically reviewed the literature that employs SFA to measure the efficiency of the HE sector at the institutional level. Our survey contains 63 studies, including 208 estimations, which were published in peer-reviewed journals until 2017.

The first part of the review thereby provides insights on how the literature developed in the last three decades (section 4.1). We confirmed the general perception that an increasing number of studies are evaluating the efficiency of HE institutions. Most studies are written by natives, using panel data provided by the respective statistical office to assess public universities. However, the utilized samples, in particular, the number of evaluated institutions and the evaluated timeframe, vary greatly between studies. Surprisingly, few authors compare efficiency across countries. Since suitable datasets are gradually emerging, we expect cross-country evaluations to become more frequent in the future.

Subsequent to this first insight, we turned toward the methodical details and evaluated the chosen underlying assumptions (section 4.2) and specifications (section 4.3). Evidently, a majority of studies chose to evaluate translog cost-functions, assuming a half-normal distribution of the

[31] Table A.1 in the appendix contains all the reviewed studies. If an attentive reader is aware of an additional study fitting the criteria for inclusion and exclusion (explained in section 3.1), we would be grateful for a respective notice.

efficiency term. The specifications proposed by Battese and Coelli (1995), including determinants of efficiency, and by Battese and Coelli (1992), for the estimation of time-variant efficiency, are the two most frequently used approaches in the evaluation of HE institutions. However, the most prominent development concerns the consideration of heterogeneity, with a large variety of applied specifications. A further advancement in the SFA literature concerns the distinction between short- and long-term efficiency, which provides additional information and allows for a more accurate measurement of heterogeneity. Given its advantages, we believe the latter specification to be a rewarding path for future studies. A further promising step will be the inclusion of determinants to explain short- and long-term efficiency. The study by Badunenko and Kumbhakar (2017), taking this approach for the evaluation of banks, can thereby serve as an example. However, for the application to HE institutions, an introductory discussion would be vital, arguing why the respective determinant is suited for one or the other efficiency term. However, given that the literature has had difficulties identifying consistent determinants so far, this examination will be challenging. In addition, the review pointed to specifications aiming to avoid the usually necessary distribution assumptions, as another interesting and not yet exhausted topic of interest.

Following the methodical details, we reviewed the employed factors, including inputs, outputs, prices and dummy variables that were used in the frontier literature (section 4.4). The evaluation exposes that there seems to be a core composition of input and output variables that represent the expenditures as well as the teaching and research tasks of universities. However, the choice of variables complementing those is diverse. More emphasis could be laid on the third mission activities of institutions. Surprisingly, few authors vary their composition of factors, especially the inputs, within one study. Unfortunately, this limits the possibility to debate the impact of the chosen factors on the efficiency level further. Studies that do vary the employed variables show that the efficiency values differ slightly with the composition. We, therefore, recommend that future researchers put more emphasis on the robustness of their chosen variables, comparing different compositions of factors. Based on the review, it can be expected that the inclusion of input prices will become customary. While dummy variables are already frequently included in efficiency evaluations, a clear line of reasoning for doing so is often missing.

The succeeding assessment of employed determinants of efficiency (section 4.5) shows that a wide variety of factors is used, but that there is no visible standard emerging. While variables concerning the teaching side of universities are employed quite frequently, through students or quality related variables, only a few studies include further determinants depicting the research aspect of institutions. In contrast, variables representing institutional characteristics are more popular, such as the personal structure, the budget, the subject specialization and the region in which the institution is located. However, the choice of variables seems often to be data driven. Some of the chosen variables are (arguably) not suited as determinants of efficiency, since they are within the control of the institutions. Hence, the determinants of efficiency seem to have great potential, but their implementation contains many inconsistencies so far. We, therefore, think that a debate on the choice and suitability of determinants is necessary and a resulting clear recommendation would be truly beneficial for future research.

In addition to the above noted possibilities for extensions, we would like to focus attention on some further limitations and a promising current developments. Two major concerns that we pointed out in the discussion relate to the sometimes omitted information and irregular presentation of efficiency values. Since both are avoidable shortcomings, we hope that future authors will ensure that all necessary information is provided in a structured manner. In addition, we think that the further discussion on the representation of research activities of universities would be valuable. In light of the prominent critique concerning the usage of research grants as an indicator, it is surprising that the literature has so far refrained from introducing adjustments to this variable. Likewise, almost no study in our sample considers determinants of efficiency that reflect the research characteristics of institutions. Beyond these shortcomings, we think that the paper by Johnes and Tsionas (2017) forms a great basis for future studies. The authors look at the dynamics of inefficiency, focusing on mergers of English HE institutions. Given that the majority of HE sectors in Europe have undergone reforms in the last decades, the evaluation of their success would be rewarding.

While we hope that our review of the literature is useful to the research community, we want to warn against taking the easy route and using this review to simply verify that the already chosen factors are valid. Rather, we hope that our review will encourage future research and discussion on the above-shown possibilities and shortcomings.

References

Abbott, Malcolm, and Chris Doucouliagos. 2009. "Competition and Efficiency: Overseas Students and Technical Efficiency in Australian and New Zealand Universities." *Education Economics* 17 (1): 31–57.

Agasisti, Tommaso. 2016. "Cost Structure, Productivity and Efficiency of the Italian Public Higher Education Industry 2001–2011." *International Review of Applied Economics* 30 (1): 48–68.

Agasisti, Tommaso, Cristian Barra, and Roberto Zotti. 2016. "Evaluating the Efficiency of Italian Public Universities (2008–2011) in Presence of (Unobserved) Heterogeneity." *Socio-Economic Planning Sciences* 55: 47–58.

Agasisti, Tommaso, and Clive Belfield. 2017. "Efficiency in the Community College Sector: Stochastic Frontier Analysis." *Tertiary Education and Management* 23 (3): 237–259.

Agasisti, Tommaso, and Sabine Gralka. 2017. "The Transient and Persistent Efficiency of Italian and German Universities: A Stochastic Frontier Analysis." *CEPIE Working Paper* 14/17.

Agasisti, Tommaso, and Carla Haelermans. 2016. "Comparing Efficiency of Public Universities among European Countries: Different Incentives Lead to Different Performances." *Higher Education Quarterly* 70 (1): 81–104.

Agasisti, Tommaso, and Geraint Johnes. 2009. "Beyond Frontiers: Comparing the Efficiency of Higher Education Decision-Making Units across More than One Country." *Education Economics* 17 (1): 59–79.

Agasisti, Tommaso, and Geraint Johnes. 2010. "Heterogeneity and the Evaluation of Efficiency: The Case of Italian Universities." *Applied Economics* 42 (11): 1365–1375.

Agasisti, Tommaso, and Geraint Johnes. 2015. "Efficiency, Costs, Rankings and Heterogeneity: The Case of US Higher Education." *Studies in Higher Education* 40 (1): 60–82.

Aiello, Francesco, and Graziella Bonanno. 2016. "Efficiency in Banking: A Meta-Regression Analysis." *International Review of Applied Economics* 30 (1): 112–149.

Aigner, Dennis, CA Knox Lovell, and Peter Schmidt. 1977. "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics* 6 (1): 21–37.

Bachan, Ray. 2017. "Grade Inflation in UK Higher Education." *Studies in Higher Education* 42 (8): 1580–1600.

Badunenko, Oleg, and Subal C. Kumbhakar. 2016. "When, Where and How to Estimate Persistent and Transient Efficiency in Stochastic Frontier Panel Data Models." *European Journal of Operational Research* 225 (1): 272–287.

Badunenko, Oleg, and Subal C. Kumbhakar. 2017. "Economies of Scale, Technical Change and Persistent and Time-Varying Cost Efficiency in Indian Banking: Do Ownership, Regulation and Heterogeneity Matter?" *European Journal of Operational Research* 260 (2): 789–803.

Barra, Cristian, Raffaele Lagravinese, and Roberto Zotti. 2018. "Does Econometric Methodology Matter to Rank Universities? An Analysis of Italian Higher Education System." *Socio-Economic Planning Sciences* 62: 104–120.

Battese, George E., and Tim J. Coelli. 1988. "Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data." *Journal of Econometrics* 38 (3): 387–399.

Battese, George E., and Tim J. Coelli. 1992. *Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India*. Springer.

Battese, George E., and Tim J. Coelli. 1995. "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data." *Empirical Economics* 20 (2): 325–332.

Battese, George E., DS Prasada Rao, and Christopher J. O'Donnell. 2004. "A Metafrontier Production Function for Estimation of Technical Efficiencies and Technology Gaps for Firms Operating under Different Technologies." *Journal of Productivity Analysis* 21 (1): 91–103.

Baumol, William J., John C. Panzar, Robert D. Willig, Elizabeth E. Bailey, Dietrich Fischer, and Dietrich Fischer. 1982. *Contestable Markets and the Theory of Industry Structure*. Harcourt Brace Jovanovich, Inc.

Bayraktar, Erkan, Ekrem Tatoglu, and Selim Zaim. 2013. "Measuring the Relative Efficiency of Quality Management Practices in Turkish Public and Private Universities." *Journal of the Operational Research Society* 64 (12): 1810–1830.

Bolli, Thomas, Maria Olivares, Andrea Bonaccorsi, Cinzia Daraio, Adela Garcia Aracil, and Benedetto Lepori. 2016. "The Differential Effects of Competitive Funding on the Production Frontier and the Efficiency of Universities." *Economics of Education Review* 52: 91–104.

Booth, Andrew, Anthea Sutton, and Diana Papaioannou. 2016. *Systematic Approaches to a Successful Literature Review*. Sage.

Bravo-Ureta, Boris E., Daniel Solís, Víctor H. Moreira López, José F. Maripani, Abdourahmane Thiam, and Teodoro Rivas. 2007. "Technical Efficiency in Farming: A Meta-Regression Analysis." *Journal of Productivity Analysis* 27 (1): 57–72.

Campbell, Randall, Kevin Rogers, and Jon Rezek. 2008. "Efficient Frontier Estimation: A Maximum Entropy Approach." *Journal of Productivity Analysis* 30 (3): 213–221.

Canton, Erik, Anna Thum-Thysen, and Peter Voigt. 2018. "Economists' Musings on Human Capital Investment: How Efficient Is Public Spending on Education in EU Member States?" *European Commission Discussion Paper* 81: 1–30.

Cardoso, Catarina Moura e Sa, and Geetha Ravishankar. 2015. "Productivity Growth and Convergence: A Stochastic Frontier Analysis." *Journal of Economic Studies* 42 (2): 224–236.

Castano, Mary Caroline N., and Emilyn C. Cabanda. 2007. "Performance Evaluation of the Efficiency of Philippine Private Higher Educational Institutions: Application of Frontier Approaches." *International Transactions in Operational Research* 14 (5): 431–444.

Caudill, Steven B. 2003. "Estimating a Mixture of Stochastic Frontier Regression Models via the EM Algorithm: A Multiproduct Cost Function Application." *Empirical Economics* 28 (3): 581–598.

Chapple, Wendy, Andy Lockett, Donald Siegel, and Mike Wright. 2005. "Assessing the Relative Performance of UK University Technology Transfer Offices: Parametric and Non-Parametric Evidence." *Research Policy* 34 (3): 369–384.

Coelli, Timothy J., Dodla Sai Prasada Rao, Christopher J. O'Donnell, and George Edward Battese. 2005. *An Introduction to Efficiency and Productivity Analysis*. Springer Science & Business Media.

Colombi, Roberto, Subal C. Kumbhakar, Gianmaria Martini, and Giorgio Vittadini. 2014. “Closed-Skew Normality in Stochastic Frontiers with Individual Effects and Long/Short-Run Efficiency.” *Journal of Productivity Analysis* 42 (2): 123–136.

Cornwell, Christopher, Peter Schmidt, and Robin C. Sickles. 1990. “Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels.” *Journal of Econometrics* 46 (1–2): 185–200.

Daghbashyan, Zara, Enrico Deiaco, and Maureen McKelvey. 2014. *How and Why Does Cost Efficiency of Universities Differ across European Countries? An Explorative Attempt Using New Microdata*. Edward Elgar Cheltenham.

Das, Subhrabaran, and Monalisa Das. 2014. “Technical Efficiency and Performance of the Higher Educational Institutions: A Study of Affiliated Degree Colleges of Barak Valley in Assam.” *Journal of Social and Economic Policy* 11 (1): 59–70.

De Witte, Kristof, and Laura López-Torres. 2017. “Efficiency in Education: A Review of Literature and a Way Forward.” *Journal of the Operational Research Society* 68 (4): 339–363.

Eagan, Kevin, and Marvin A. Titus. 2016. “Examining Production Efficiency in Higher Education: The Utility of Stochastic Frontier Analysis.” In *Higher Education: Handbook of Theory and Research*. Springer.

Emrouznejad, Ali, Barnett R. Parker, and Gabriel Tavares. 2008. “Evaluation of Research in Efficiency and Productivity: A Survey and Analysis of the First 30 Years of Scholarly Literature in DEA.” *Socio-Economic Planning Sciences* 42 (3): 151–157.

Erkoc, Taptuk Emre. 2015. “Assessing the Research Performance in Higher Education with Stochastic Distance Function Approach.” *International Journal of Education Economics and Development* 6 (4): 366–380.

EUMIDA. 2009. *Final Study Report: Feasibility Study for Creating a European University Data Collection. European Commission*. Research Directorate-General C-European Research Area Universities and Researchers. Contract No. RTD/C/C4/2009/0233402.

Farsi, Mehdi, Massimo Filippini, and William Greene. 2006. “Application of Panel Data Models in Benchmarking Analysis of the Electricity Distribution Sector.” *Annals of Public and Cooperative Economics* 77 (3): 271–290.

Fieger, Peter, Renato Villano, and Ray Cooksey. 2016. “Efficiency of Australian Technical and Further Education Providers.” *International Journal of Training Research* 14 (1): 62–75.

Filippini, Massimo, and William Greene. 2016. “Persistent and Transient Productive Inefficiency: A Maximum Simulated Likelihood Approach.” *Journal of Productivity Analysis* 45 (2): 187–196.

Fried, Harold O., CA Knox Lovell, Shelton S. Schmidt, and Shelton S. Schmidt. 2008. *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press.

Fu, Tsu-Tan, Cliff J. Huang, and Flora F. Tien. 2008. “University Cost Structure in Taiwan.” *Contemporary Economic Policy* 26 (4): 651–662.

Gralka, Sabine. 2018. “Persistent Inefficiency in the Higher Education Sector: Evidence from Germany.” *Education Economics* 26 (4): 373–392.

Gralka, Sabine, Klaus Wohlrabe, and Lutz Bornmann. 2018. “How to Measure Research Efficiency in Higher Education? Research Grants vs. Publication Output.” *CESifo Working Paper Series 7055*.

Greene, William H. 2005. “Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model.” *Journal of Econometrics* 126 (2): 269–303.

Greene, William H. 2008. “The Econometric Approach to Efficiency Analysis.” *The Measurement of Productive Efficiency and Productivity Growth*, 92–250.

Horne, Jocelyn, and Baiding Hu. 2008. “Estimation of Cost Efficiency of Australian Universities.” *Mathematics and Computers in Simulation* 78 (2–3): 266–275.

Hu, Jin-Li, Chih-Hai Yang, and Chiang-Ping Chen. 2014. “R&D Efficiency and the National Innovation System: An International Comparison Using the Distance Function Approach.” *Bulletin of Economic Research* 66 (1): 55–71.

Izadi, Hooshang, Geraint Johnes, Reza Oskrochi, and Robert Crouchley. 2002. “Stochastic Frontier Estimation of a CES Cost Function: The Case of Higher Education in Britain.” *Economics of Education Review* 21 (1): 63–71.

Johnes, Geraint. 1998. “The Costs of Multi-Product Organizations and the Heuristic Evaluation of Industrial Structure.” *Socio-Economic Planning Sciences* 32 (3): 199–209.

Johnes, Geraint. 2004. “A Fourth Desideratum: The CES Cost Function and the Sustainable Configuration of Multiproduct Firms.” *Bulletin of Economic Research* 56 (4): 329–332.

Johnes, Geraint, Ana S. Camanho, and Maria S. Portela. 2008. “Assessing Efficiency of Portuguese Universities through Parametric and Non-Parametric Methods.” *Portuguese Journal of Management Studies* 13 (1): 39–66.

Johnes, Geraint, and Jill Johnes. 1993. “Measuring the Research Performance of UK Economics Departments: An Application of Data Envelopment Analysis.” *Oxford Economic Papers*, 332–347.

Johnes, Geraint, and Jill Johnes. 2009. “Higher Education Institutions’ Costs and Efficiency: Taking the Decomposition a Further Step.” *Economics of Education Review* 28 (1): 107–113.

Johnes, Geraint, and Jill Johnes. 2016. “Costs, Efficiency, and Economies of Scale and Scope in the English Higher Education Sector.” *Oxford Review of Economic Policy* 32 (4): 596–614.

Johnes, Geraint, Jill Johnes, and Emmanuel Thanassoulis. 2008. “An Analysis of Costs in Institutions of Higher Education in England.” *Studies in Higher Education* 33 (5): 527–549.

Johnes, Geraint, Jill Johnes, Emmanuel Thanassoulis, Pam Lenton, and Ali Emrouznejad. 2005. *An Exploratory Analysis of the Cost Structure of Higher Education in England*. UK Department for Education and Skills.

Johnes, Geraint, and Manuel Salas-Velasco. 2007. “The Determinants of Costs and Efficiencies Where Producers Are Heterogeneous: The Case of Spanish Universities.” *Economics Bulletin* 4 (15): 1–9.

Johnes, Geraint, and Astrid Schwarzenberger. 2011. “Differences in Cost Structure and the Evaluation of Efficiency: The Case of German Universities.” *Education Economics* 19 (5): 487–499.

Johnes, Jill. 2006. “Data Envelopment Analysis and Its Application to the Measurement of Efficiency in Higher Education.” *Economics of Education Review* 25 (3): 273–288.

Johnes, Jill. 2014. "Efficiency and Mergers in English Higher Education 1996/97 to 2008/9: Parametric and Non-Parametric Estimation of the Multi-Input Multi-Output Distance Function." *The Manchester School* 82 (4): 465–487.

Johnes, Jill. 2015. "Operational Research in Education." *European Journal of Operational Research* 243 (3): 683–696.

Johnes, Jill, and Mike G. Tsionas. 2017. "Dynamics of Inefficiency and Merger in English Higher Education From 1996/97 to 2008/9: A Comparison of Pre-Merging, Post-Merging and Non-Merging Universities Using Bayesian Methods." *The Manchester School*, 1–37.

Jondrow, James, CA Knox Lovell, Ivan S. Materov, and Peter Schmidt. 1982. "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model." *Journal of Econometrics* 19 (2): 233–238.

Kempkes, Gerhard, and Carsten Pohl. 2008. "Do Institutions Matter for University Cost Efficiency? Evidence from Germany." *CESifo Economic Studies* 54 (2): 177–203.

Kempkes, Gerhard, and Carsten Pohl. 2010. "The Efficiency of German Universities–Some Evidence from Nonparametric and Parametric Methods." *Applied Economics* 42 (16): 2063–2079.

Kulshreshtha, Praveen, and Tapan Kumar Nayak. 2015. "Efficiency of Higher Technical Educational Institutions in India." *Archives of Business Research* 3 (1): 94–122.

Kumbhakar, Subal C. 1990. "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency." *Journal of Econometrics* 46 (1): 201–211.

Kumbhakar, Subal C., Soumendra Ghosh, and J. Thomas McGuckin. 1991. "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in US Dairy Farms." *Journal of Business & Economic Statistics* 9 (3): 279–286.

Kumbhakar, Subal C., and Almas Heshmati. 1995. "Efficiency Measurement in Swedish Dairy Farms: An Application of Rotating Panel Data, 1976–88." *American Journal of Agricultural Economics* 77 (3): 660–674.

Kumbhakar, Subal C., Gudbrand Lien, and J. Brian Hardaker. 2014. "Technical Efficiency in Competing Panel Data Models: A Study of Norwegian Grain Farming." *Journal of Productivity Analysis* 41 (2): 321–337.

Kumbhakar, Subal C., and CA Knox Lovell. 2003. *Stochastic Frontier Analysis*. Cambridge University Press.

Kumbhakar, Subal C., Hongren Wang, and Alan P. Horncastle. 2015. *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*. Cambridge University Press.

Kuo, Jenn-Shyong, and Yi-Cheng Ho. 2008. "The Cost Efficiency Impact of the University Operation Fund on Public Universities in Taiwan." *Economics of Education Review* 27 (5): 603–612.

Laureti, Tiziana. 2008. "Modelling Exogenous Variables in Human Capital Formation through a Heteroscedastic Stochastic Frontier." *International Advances in Economic Research* 14 (1): 76–89.

Laureti, Tiziana, Luca Secondi, and Luigi Biggeri. 2014. "Measuring the Efficiency of Teaching Activities in Italian Universities: An Information Theoretic Approach." *Economics of Education Review* 42: 147–164.

Lenton, Pamela. 2008. "The Cost Structure of Higher Education in Further Education Colleges in England." *Economics of Education Review* 27 (4): 471–482.

Lu, Yung-Hsiang, and Ku-Hsieh Chen. 2013. "Appraising the Cost Efficiency of Higher Technological and Vocational Education Institutions in Taiwan Using the Metafrontier Cost-Function Model." *Research in Higher Education* 54 (6): 627–663.

Mamun, Shamsul Arifeen Khan. 2011. "Are Public Universities of Bangladesh Cost Efficient? An Empirical Evidence." *South Asia Economic Journal* 12 (2): 221–237.

McMillan, Melville L., and Wing H. Chan. 2006. "University Efficiency: A Comparison and Consolidation of Results from Stochastic and Non-Stochastic Methods." *Education Economics* 14 (1): 1–30.

Mensah, Yaw M., and Robert Werner. 2003. "Cost Efficiency and Financial Flexibility in Institutions of Higher Education." *Journal of Accounting and Public Policy* 22 (4): 293–323.

Miranda, Rodrigo, Maria Cristina Gramani, and Eduardo Andrade. 2012. "Technical Efficiency of Business Administration Courses: A Simultaneous Analysis Using DEA and SFA." *International Transactions in Operational Research* 19 (6): 847–862.

Mutz, Rüdiger, Lutz Bornmann, and Hans-Dieter Daniel. 2017. "Are There Any Frontiers of Research Performance? Efficiency Measurement of Funded Research Projects with the Bayesian Stochastic Frontier Analysis for Count Data." *Journal of Informetrics* 11 (3): 613–628.

Nemoto, Jiro, and Noriko Furumatsu. 2014. "Scale and Scope Economies of Japanese Private Universities Revisited with an Input Distance Function Approach." *Journal of Productivity Analysis* 41 (2): 213–226.

Neyman, Jerzy, and Elizabeth L. Scott. 1948. "Consistent Estimates Based on Partially Consistent Observations." *Econometrica: Journal of the Econometric Society*, 1–32.

Olivares, Maria, and Heike Wetzel. 2014. "Competing in the Higher Education Market: Empirical Evidence for Economies of Scale and Scope in German Higher Education Institutions." *CESifo Economic Studies* 60 (4): 653–680.

Orea, Luis, and Subal C. Kumbhakar. 2004. "Efficiency Measurement Using a Latent Class Stochastic Frontier Model." *Empirical Economics* 29 (1): 169–183.

Pitt, Mark M., and Lung-Fei Lee. 1981. "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry." *Journal of Development Economics* 9 (1): 43–64.

Rhaiem, Mehdi. 2017. "Measurement and Determinants of Academic Research Efficiency: A Systematic Review of the Evidence." *Scientometrics* 110 (2): 581–615.

Robst, John. 2001. "Cost Efficiency in Public Higher Education Institutions." *The Journal of Higher Education* 72 (6): 730–750.

Rusnák, Marek, Tomas Havranek, and Roman Horváth. 2013. "How to Solve the Price Puzzle? A Meta-Analysis." *Journal of Money, Credit and Banking* 45 (1): 37–70.

Rzadzinski, Leslaw, and Anna Sworowska. 2016. "Parametric and Non Parametric Methods for Efficiency Assessment of State Higher Vocational Schools in 2009-2011." *Entrepreneurial Business and Economics Review* 4 (1): 95–112.

Sav, G. Thomas. 2011. "Cost Efficiencies and Rankings of Flagship Universities." *American Journal of Economics and Business Administration* 3 (4): 596–603.

Sav, G. Thomas. 2012a. "Efficiency Estimates and Rankings Employing Data Envelopment and Stochastic Frontier Analyses: Evaluating the Management of U.S. Public Colleges." *Information Management and Business Review*, 444–452.

Sav, G. Thomas. 2012b. "Frontier and Envelopment Evaluations of University Graduation Efficiencies and Productivities: Elements for Performance Based Funding." *Problems and Perspectives in Management*, 71–79.

Sav, G. Thomas. 2012c. "Cost Inefficiencies and Rankings of Ivy Universities: Stochastic Panel Estimates." *International Journal of Economics and Finance* 4 (3): 3–12.

Sav, G. Thomas. 2012d. "Stochastic Cost Inefficiency Estimates and Rankings of Public and Private Research and Doctoral Granting Universities." *Journal of Knowledge Management, Economics, and Information Technology* 2 (3): 1–20.

Sav, G. Thomas. 2012e. "Managing Operating Efficiencies of Publicly Owned Universities: American University Stochastic Frontier Estimates Using Panel Data." *Advances in Management and Applied Economics* 2 (1): 1–23.

Sav, G. Thomas. 2012f. "Minority Serving College and University Cost Efficiencies." *Journal of Social Sciences*, 54–60.

Sav, G. Thomas. 2012g. "For-Profit College Entry and Cost Efficiency: Stochastic Frontier Estimates vs Two-Year Public and Non-Profit Colleges." *International Business Research* 5 (3): 26–32.

Sav, G. Thomas. 2012h. "Is the Production of Religious Knowledge Efficient? Managing Faith Related Postsecondary Institutions." *Journal of Knowledge Management, Economics, and Information Technology* 2 (1): 1–13.

Sav, G. Thomas. 2012i. "Stochastic Cost Frontier and Inefficiency Estimates of Public and Private Universities: Does Government Matter?" *International Advances in Economic Research* 18 (2): 187–198.

Sav, G. Thomas. 2012j. "Historically Black College and University Operating Cost Efficiencies: Stochastic Cost Estimates and Comparisons to Predominately White Institutions." *International Journal of Education Economics and Development* 3 (3): 252–263.

Sav, G. Thomas. 2012k. "Does Faculty Tenure Improve Student Graduation Rates?" *Journal of Business Management and Applied Economics* 1 (5): 1–19.

Sav, G. Thomas. 2012l. "Female Faculty, Tenure, and Student Graduation Success: Efficiency Implications for University Funding." *International Journal of Business Management & Economic Research* 3 (6): 633–640.

Sav, G. Thomas. 2013. "Private Philanthropy in Financing Public Universities: Fundraising Stochastic Frontier and Efficiency Evaluations." *International Research Journal of Finance and Economics* 103: 1–13.

Sav, G. Thomas. 2016. "Declining State Funding and Efficiency Effects on Public Higher Education: Government Really Does Matter." *International Advances in Economic Research* 22 (4): 397–408.

Schmidt, Peter, and Robin C. Sickles. 1984. "Production Frontiers and Panel Data." *Journal of Business & Economic Statistics* 2 (4): 367–374.

Sriboonchitta, Songsak. 2012. "Evaluation of Cost Efficiency of Thai Public Universities." *International Journal of Intelligent Technologies and Applied Statistics* 5 (4): 361–374.

Stevens, Philip Andrew. 2005. "A Stochastic Frontier Analysis of English and Welsh Universities." *Education Economics* 13 (4): 355–374.

Titus, Marvin A. 2009. "Bachelor's Degree Productivity X-Inefficiency: The Role of State Higher Education Policy." *Journal of College Student Retention: Research, Theory & Practice* 11 (1): 7–32.

Titus, Marvin A., and Brian Pusser. 2011. "States' Potential Enrollment of Adult Students: A Stochastic Frontier Analysis." *Research in Higher Education* 52 (6): 555–571.

Titus, Marvin A., Adriana Vamosiu, and Kevin R. McClure. 2016. "Are Public Master's Institutions Cost Efficient? A Stochastic Frontier and Spatial Analysis." *Research in Higher Education* 58 (5): 469–496.

Tsionas, Efthymios G., and Subal C. Kumbhakar. 2014. "Firm Heterogeneity, Persistent And Transient Technical Inefficiency: A Generalized True Random-Effects Model." *Journal of Applied Econometrics* 29 (1): 110–132.

Wang, Hung-Jen, and Chia-Wen Ho. 2010. "Estimating Fixed-Effect Panel Stochastic Frontier Models by Model Transformation." *Journal of Econometrics* 157 (2): 286–296.

Wang, Hung-Jen, and Peter Schmidt. 2002. "One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels." *Journal of Productivity Analysis* 18 (2): 129–144.

Warning, Susanne. 2007. *The Economic Analysis of Universities: Strategic Groups and Positioning*. Cheltenham UK: Elgar.

Worthington, Andrew C. 2001. "An Empirical Survey of Frontier Efficiency Measurement Techniques in Education." *Education Economics* 9 (3): 245–268.

Worthington, Andrew C., and Helen Higgs. 2011. "Economies of Scale and Scope in Australian Higher Education." *Higher Education* 61 (4): 387–414.

Zhang, Liang, Wei Bao, and Liang Sun. 2016. "Resources and Research Production in Higher Education: A Longitudinal Analysis of Chinese Universities, 2000–2010." *Research in Higher Education* 57 (7): 869–891.

Zoghbi, Ana Carolina, Fabiana Rocha, and Enlinson Mattos. 2013. "Education Production Efficiency: Evidence from Brazilian Universities." *Economic Modelling* 31: 94–103.

Appendix

Table 1: Considered publications, sorted by year

Reference	Title	Country	Type of Function
Johnes (1998)	The Costs of Multi-product Organizations and the Heuristic Evaluation of Industrial Structure	United Kingdom	Cost function
Robst (2001)	Cost Efficiency in Public Higher Education	United States	Cost function
Izadi et al. (2002)	Stochastic frontier estimation of a CES cost function: the case of higher education in Britain	United Kingdom	Cost function
Mensah and Werner (2003)	Cost efficiency and financial flexibility in institutions of higher education	United States	Cost function
Chapple et al. (2005)	Assessing the relative performance of U.K. university technology transfer offices: parametric and non-parametric evidence	United Kingdom	Production function
Stevens (2005)	A stochastic frontier analysis of English and Welsh universities	United Kingdom	Cost function
McMillan and Chan (2006)	University Efficiency: A Comparison and Consolidation of Results from Stochastic and Non-stochastic Methods	Canada	Cost function
Castano and Cabanda (2007)	Performance evaluation of the efficiency of Philippine Private Higher Educational Institutions: application of frontier approaches	Philippines	Production function
Johnes and Salas-Velasco (2007)	The determinants of costs and efficiencies where producers are heterogeneous: The case of Spanish universities	Spain	Cost function
Johnes, Camanho and Portela (2008)	Assessing efficiency of Portuguese universities through parametric and non-parametric methods	Portugal	Cost function
Fu, Huang and Tien (2008)	University Cost Structure in Taiwan	Taiwan	Cost function
Horne and Hu (2008)	Estimation of cost efficiency of Australian universities	Australia	Cost function
Johnes, Johnes, and Thanassoulis (2008)	An analysis of costs in institutions of higher education in England	United Kingdom	Cost function
Kempkes and Pohl (2008)	Do Institutions Matter for University Cost Efficiency? Evidence from Germany	Germany	Cost function
Kuo and Ho (2008)	The cost efficiency impact of the university operation fund on public universities in Taiwan	Taiwan	Cost function
Lenton (2008)	The cost structure of higher education in further education colleges in England	United Kingdom, United States	Cost function

Abbott and Doucouliagos (2009)	Competition and efficiency: overseas students and technical efficiency in Australian and New Zealand universities	Australia, New Zealand	Distance function
Johnes and Johnes (2009)	Higher education institutions' costs and efficiency: Taking decomposition a further step	United Kingdom	Cost function
Agasisti and Johnes (2009)	Heterogeneity and the evaluation of efficiency: the case of Italian universities	Italy	Cost function
Kempkes and Pohl (2010)	The efficiency of German universities - some evidence from nonparametric and parametric methods	Germany	Cost function
Johnes and Schwarzenberger (2011)	Differences in cost structure and the evaluation of efficiency: the case of German universities	Germany	Cost function
Mamun (2011)	Are Public Universities of Bangladesh Cost Efficient? An Empirical Evidence	Bangladesh	Cost function
Sav (2011)	Cost Efficiencies and Rankings of Flagship Universities	United States	Cost function
Worthington and Higgs (2011)	Economies of scale and scope in Australian higher education	Australia	Cost function
Miranda, Gramani and Andrade (2012)	Technical efficiency of business administration courses: a simultaneous analysis using DEA and SFA	Brazil	Production function
Sav (2012a)	Is the Production of Religious Knowledge Efficient? Managing Faith Related Postsecondary Institutions	United States	Cost function
Sav (2012b)	Managing Operating Efficiencies of Publicly Owned Universities: American University Stochastic Frontier Estimates Using Panel Data	United States	Cost function
Sav (2012c)	Does Faculty Tenure Improve Student Graduation Rates?	United States	Production function
Sav (2012d)	Female Faculty, Tenure, and Student Graduation Success: Efficiency Implications for University Funding	United States	Production function
Sav (2012e)	Frontier and envelopment evaluations of university graduation efficiencies and productivities: elements of performance based funding	United States	Production function
Sav (2012f)	Stochastic Cost Inefficiency Estimates and Rankings of Public and Private Research and Doctoral Granting Universities	United States	Cost function
Sav (2012g)	For-Profit College Entry and Cost Efficiency: Stochastic Frontier Estimates vs Two-Year Public and Non-Profit Colleges	United States	Cost function
Sav (2012h)	Cost Inefficiencies and Rankings of Ivy Universities: Stochastic Panel Estimates	United States	Cost function

Sav (2012i)	Minority Serving College and University Cost Efficiencies	United States	Cost function
Sav (2012j)	Historically black college and university operating cost efficiencies: stochastic cost estimates and comparisons to predominately white institutions	United States	Cost function
Sav (2012k)	Stochastic Cost Frontier and Inefficiency Estimates of Public and Private Universities: Does Government Matter?	United States	Cost function
Sav (2012l)	Efficiency Estimates and Rankings Employing Data Envelopment and Stochastic Frontier Analyses: Evaluating the Management of U.S. Public Colleges	United States	Production function
Sriboonchitta (2012)	Evaluation of Cost Efficiency of Thai Public Universities	Thailand	Cost function
Lu and Chen (2013)	Appraising the cost efficiency of higher technological and vocational education institutions in Taiwan using the metafrontier cost-function model	Taiwan	Cost function
Zoghbi, Rocha and Mattos (2013)	Education production efficiency: Evidence from Brazilian universities	Brazil	Production function
Sav (2013)	Private Philanthropy in Financing Public Universities: Fundraising Stochastic Frontier and Efficiency Evaluations	United States	Production function
Das and Das (2014)	Technical Efficiency and Performance of the Higher Educational Institutions: A Study of Affiliated Degree Colleges of Barak Valley in Assam	India	Production function
Johnes (2014)	Efficiency and Mergers in English Higher Education 1996/97 to 2008/9: Parametric and Non-parametric Estimation of the Multi-input Multi-output Distance Function	United Kingdom	Distance function
Laureti, Secondi, and Biggeri (2014)	Measuring the efficiency of teaching activities in Italian universities: An information theoretic approach	Italy	Production function
Nemoto and Furumatsu (2014)	Scale and scope economies of Japanese private universities revisited with an input distance function	Japan	Distance function
Olivares and Wetzel (2014)	Competing in the Higher Education Market: Empirical Evidence for Economies of Scale and Scope in German Higher Education Institutions	Germany	Distance function
Agasisti and Johnes (2015)	Efficiency, costs, rankings and heterogeneity: the case of US higher education	United States	Cost function
Erkoc (2015)	Assessing the research performance in higher education with stochastic distance function approach	Turkey	Distance function

Kulshreshtha and Nayak (2015)	Efficiency of Higher Technical Educational Institutions in India	India	Distance function
Agasisti and Haelermans (2016)	Comparing efficiency of public universities among European countries: Different incentives lead to different performances.	Italy, Netherlands	Cost function
Agasisti (2016)	Cost structure, productivity and efficiency of the Italian public higher education industry 2001–2011	Italy	Cost function
Agasisti, Barra and Zotti (2016)	Evaluating the efficiency of Italian public universities (2008-2011) in presence of (unobserved) heterogeneity	Italy	Distance function
Bolli et al. (2016)	The differential effects of competitive funding on the production frontier and the efficiency of universities	Finland, Italy, Netherlands, Norway, Portugal, Spain, Switzerland, United Kingdom	Distance function
Fieger, Villano and Cooksey (2016)	Efficiency of Australian technical and further education providers	Australia	Production function
Johnes and Johnes (2016)	Cost, efficiency, and economies of scale and scope in the English higher education sector	United Kingdom	Cost function
Rzadzinski and Sworowska (2016)	Parametric and Non-Parametric Methods for Efficiency Assessment of State Higher Vocational Schools in 2009-2011	Poland	Production function
Sav (2016)	Declining State Funding and Efficiency Effects on Public Higher Education: Government Really Does Matter	United States	Cost function
Titus, Vamosiu, and McClure (2016)	Are Public Masters Institutions Cost Efficient? A Stochastic Frontier and Spatial Analysis	United States	Distance function
Zhang, Bao and Sun (2016)	Resources and Research Production in Higher Education: A Longitudinal Analysis of Chinese Universities, 2000-2010	China	Production function
Agasisti and Belfield (2017)	Efficiency in the community college sector: stochastic frontier analysis	United States	Production function
Bachan (2017)	Grade inflation in UK higher education	United Kingdom	Production function
Barra, Lagravinese and Zotti (2018)	Does econometric methodology matter to rank universities? An analysis of Italian higher education system	Italy	Distance function
Gralka (2018)	Persistent inefficiency in the higher education sector: evidence from Germany	Germany	Cost function